ON THE REPRESENTATION OF A NUMBER AS THE SUM OF ANY NUMBER OF SQUARES, AND IN PARTICULAR OF FIVE*

BY

G. H. HARDY

1. Introduction

1. 1. In a short note published recently in the Proceedings of the National Academy of Sciences[1] I sketched the outlines of a new solution of one of the most interesting and difficult problems in the Theory of Numbers, that of determining the number of representations of a given integer as the sum of five or seven squares. The method which I use is one of great power and generality, and has been applied by Mr. J. E. Littlewood, Mr. S. Ramanujan, and myself to the solution of a number of different problems; and it is probable that, in our previous writings on the subject,[2] we have explained sufficiently the general ideas on which it rests. I may therefore confine myself, for the most part, to filling in the details of my previous work. I should observe, however, that the method by which I now sum the "singular series", which plays a dominant rôle in the analysis,

* Presented to the Society, February, 1920.

[1] G. H. Hardy, *On the expression of a number as the sum of any number of squares, and in particular of five or seven*, Proceedings of the National Academy of Sciences, vol. 4 (1918), pp. 189–193.

[2] G. H. Hardy and S. Ramanujan: (1) *Une formule asymptotique pour le nombre des partitions de n*, Comptes Rendus, 2 Jan. 1917; (2) *Asymptotic formulae in Combinatory Analysis*, Proceedings of the London Mathematical Society, ser. 2, vol. 17 (1918), pp. 75–115; (3) *On the coefficients in the expansions of certain modular functions*, Proceedings of the Royal Society, (A), vol. 95 (1918), pp. 144–155:

S. Ramanujan, *On certain trigonometrical sums and their applications in the theory of numbers*, Transactions of the Cambridge Philosophical Society, vol. 22 (1918), pp. 259–276:

G. H. Hardy and J. E. Littlewood: (1) *A new solution of Waring's Problem*, Quarterly Journal of Mathematics, vol. 48 (1919), pp. 272–293: (2) *Note on Messrs. Shah and Wilson's paper entitled On an empirical formula connected with Goldbach's Theorem*, Proceedings of the Cambridge Philosophical Society, vol. 19 (1919), pp. 245–254: (3) *Some problems of 'Partitio Numerorum'*, (I) *A new solution of Waring's Problem*, Göttinger Nachrichten, 1920: (4) *Some problems of 'Partitio Numerorum'*, (II) *Proof that every large number is the sum of at most 21 biquadrates*, Mathematische Zeitschrift, 1920.

The two last papers will be published shortly.

is quite different from that which I sketched in my former note. The new method has important applications to a whole series of problems in Combinatory Analysis, concerning the representation of numbers by sums of squares, cubes, kth powers, or primes. It is in the present problem that it finds its simplest and most elegant application, and it is most instructive to work this application out in detail.

It is well known that the solution of the problem is a good deal simpler when s, the number of squares in question, does not exceed 8. If s is 2, 4, 6, or 8, the number of representations may be expressed in finite form by means of the real divisors of n; if s is 3, 5, or 7, by means of quadratic residues and non-residues. If $s > 8$, other and more recondite arithmetical functions are involved. In this paper I confine myself to the cases in which $s \leqq 8$. Among these, those in which s is odd have always been regarded as notably the more difficult, and one of my principal objects has been to place them all upon the same footing. But I generally suppose $s = 5$ or $s = 8$, cases typical of the odd and even cases respectively.

In Section 2 I construct the *singular series*

$$\frac{\pi^{\frac{1}{2}s}\, n^{\frac{1}{2}s-1}}{\Gamma(\frac{1}{2}s)} \sum_1^\infty A_k,$$

where

$$A_1 = 1, \qquad A_k = \frac{1}{k^s} \sum_h (S_{h,k})^s\, e^{-2nh\pi i/k};$$

$S_{h,k}$ denoting the Gaussian sum[3]

$$\sum_{j=1}^{k} e^{j^2 h\pi i/k},$$

and the summation extending over all positive values of h less than and prime to k. The series may be written in the form

$$\frac{\pi^{\frac{1}{2}s}\, n^{\frac{1}{2}s-1}}{\Gamma(\frac{1}{2}s)}\Big[1 + 0 + \frac{2}{3^{\frac{1}{2}s}}\cos(\tfrac{2}{3}n\pi - \tfrac{1}{2}s\pi) + \frac{2^{\frac{1}{2}s+1}}{4^{\frac{1}{2}s}}\cos(\tfrac{1}{2}n\pi - \tfrac{1}{4}s\pi) + \frac{2}{5^{\frac{1}{2}s}}\{\cos\tfrac{2}{5}n\pi + \cos(\tfrac{4}{5}n\pi - s\pi)\} + 0 + \cdots\Big],$$

the zero terms corresponding to $k = 2$ and $k = 6$.

In Section 3 I show that, when $s = 8$ or $s = 5$, the sum of the singular series is in fact $r_s(n)$, the number of representations of n as a sum of s squares. The methods used are equally applicable in the cases of 3, 4, 6, or 7 squares;

[3] In my former note I denoted a typical "rational point" on the unit circle by $e^{h\pi i/k}$, and a typical Gaussian sum by

$$\sum_0^{k-1} e^{j^2 h\pi i/k}.$$

In this paper I generally use the forms involving a 2. Each notation has special advantages for particular purposes.

but the case of two squares is abnormal.[4] Throughout this section I am very deeply indebted to a paper by Mr. Mordell, published recently in the Quarterly Journal of Mathematics.[5] My proof of the identity of the functions which I call ϑ^s and Θ_s is in fact based directly on his work. It is true that Mordell considers only the case in which s is even; but his argument is applicable in principle to either case, and was applied by him to the even case only merely because, at the time when his paper was written, he had no method for the construction, when s is odd, of the essential "principal invariant" denoted by him by χ. It is the construction of this invariant by a uniform method in all cases, through the medium of the "singular series", that is my own principal contribution to the subject.

In Section 4 I show how the singular series may be transformed into a product, and give general rules for the calculation of the terms of the product. All the results of this section are independent of the hypothesis $s \leqq 8$. In Section 5 I sum the series when $s = 8$, and obtain Jacobi's well-known results. In Section 6 I consider the case $s = 5$, supposing however that n has no squared factor, so that there is no distinction between primitive and imprimitive representations; and I obtain results equivalent to those enunciated first by Eisenstein and proved later by Smith and Minkowski. In Section 7 I consider the general case, and show that the method leads to the more complete results of Smith. I conclude, in Section 8, by some remarks as to the application of the method when $s > 8$. I do not pursue this subject further because such applications belong more naturally, either to Mr. Littlewood's and my own researches in connection with Waring's problem, or to Mr. Mordell's in connection with the general theory of modular invariants.

It will be noticed that the explicit formulas for the powers of the fundamental theta-function, such as the familiar formula

$$\vartheta^8 = (1 + 2q + 2q^4 + \cdots)^8 = 1 + 16\left(\frac{1^3 q}{1+q} + \frac{2^3 q^2}{1-q^2} + \frac{3^3 q^3}{1+q^3} + \cdots\right),$$

or the new formula[6]

$$\vartheta^5 = 1 + \frac{32}{3}\left\{\sum_{1,3,5,\ldots}\frac{1}{k^2}\sum_j\sum_{m=0}^{\infty}(mk+j)^{3/2}\,q^{mk+j} - \sum_{2,4,6,\ldots}\frac{1}{k^2}\sum_j\sum_{m=0}^{\infty}(-1)^{m+\mu}(mk+j)^{3/2}\,q^{mk+j}\right\},$$

do not appear at all in my present analysis.

[4] See Mr. Ramanujan's paper quoted in footnote 2.

[5] L. J. Mordell, *On the representations of numbers as a sum of* $2r$ *squares*, Quarterly Journal of Mathematics, vol. 48 (1917), pp. 93–104. See also a later paper by the same author, *On the representations of a number as a sum of an odd number of squares*, Transactions of the Cambridge Philosophical Society, vol. 22 (1919), no. 17, pp. 361–372.

[6] This is formula (10) of my former note, where the meaning of j and μ is explained. See also p. 360 of Mr. Mordell's second paper cited above.

In the sequel I give references only to isolated results directly required for the objects of my analysis. It is more convenient to collect here some notes concerning the older memoirs dealing with the problem.

Jacobi's classical results concerning 2, 4, 6, or 8 squares are quoted by Smith on p. 307 of his *Report on the Theory of Numbers* (*Collected Papers*, vol. 1). They are contained implicitly in §§ 40–42 of the *Fundamenta Nova* (pp. 103–115).

Liouville gave formulas relating to the cases of 10 and 12 squares in a number of short notes in the second series of the Journal des mathématiques: see in particular vol. 5, p. 143; vol. 6, p. 233; vol. 9, p. 296; vol. 10, p. 1. These notes appeared between 1860 and 1865.

Later Glaisher, in a series of papers published in the Quarterly Journal, worked out systematically all cases in which s is even and between 2 and 18 inclusive. He has given a short summary of his results in a paper *On the numbers of representations of a number as a sum of* $2r$ *squares, where* $2r$ *does not exceed* 18, published in the Proceedings of the London Mathematical Society, ser. 2, vol. 5 (1907), pp. 479–490. This paper contains full references to his more detailed work.

The results for 5 squares (for numbers which have no square divisors) were stated without proof by Eisenstein on p. 368 of vol. 35 (1847) of Crelle's Journal. They were completed by Smith, who stated the general results at the end of his memoir *On the orders and genera of quadratic forms containing more than three indeterminates* (Proceedings of the Royal Society, vol. 13 (1864), pp. 199–203, and vol. 16 (1867), pp. 197–208; *Collected Papers*, vol. 1, pp. 412–417, 510–523). No detailed proofs, however, appeared before the publication of the prize memoirs of Smith (*Mémoire sur la représentation des nombres par des sommes de cinq carrés*, Mémoires présentés par divers savants à l'Académie, vol. 29, no. 1 (1887), pp. 1–72; *Collected Papers*, vol. 2, pp. 623–680) and Minkowski (*Mémoire sur la théorie des formes quadratiques à coefficients entières*, ibid., no. 2, pp. 1–178; *Gesammelte mathematische Abhandlungen*, vol. 1, pp. 3–144).

The methods for the summation of the series

$$\Sigma\left(\frac{n}{m}\right)\frac{1}{m^2},$$

which is fundamental in the five square problem, and other series of similar type, are due to Dirichlet (*Recherches sur divers applications de l'analyse infinitésimale à la théorie des nombres*, Crelle's Journal, vol. 19 (1839), pp. 324–369, and vol. 21 (1840), pp. 1–12, 134–155; *Werke*, vol. 1, pp. 411–497) and to Cauchy (*Mémoire sur la théorie des nombres*, Mémoires de

l'Académie des Sciences, vol. 17 (1840), pp. 249–768; especially Note 12, pp. 665–699).

A systematic account of the whole theory is given by Bachmann in vol. 4 of his *Zahlentheorie*. Bachmann works out the case $s = 7$ also in detail.

2. Formal construction of the singular series

2. 1. I write, as in my former note

$$(2.11)\quad f(q) = 1 + \sum_1^\infty r_s(n) q^n = (1 + 2q + 2q^4 + \cdots)^s = \{\vartheta_3(0, \tau)\}^s = \vartheta^s,$$

where $q = e^{\pi i \tau}$ and $\Im(\tau) > 0$; and I consider the behavior of this function when q tends radially to a "rational point" $e^{2h\pi i/k}$ upon the unit circle. We may suppose that $h = 0$, $k = 1$, or that k is greater than unity and h positive, less than k, and prime to k.

If (2.11)

$$q = \boldsymbol{q} e^{2h\pi i/k},$$

so that $0 \leqq \boldsymbol{q} < 1$, $\boldsymbol{q} \to 1$, we have

$$\vartheta = 1 + 2\sum_1^\infty \boldsymbol{q}^{n^2} e^{2n^2 h\pi i/k}$$

$$= 1 + 2\sum_{j=1}^{k} \sum_{l=0}^{\infty} \boldsymbol{q}^{(lk+j)^2} e^{2(lk+j)^2 h\pi i/k}$$

$$= 1 + 2\sum_{j=1}^{k} e^{2j^2 h\pi i/k} \sum_{l=0}^{\infty} \boldsymbol{q}^{(lk+j)^2}.$$

Now

$$\sum_{k=0}^{\infty} \boldsymbol{q}^{(lk+j)^2} \sim \frac{\sqrt{\pi}}{2k}\left(\log\frac{1}{\boldsymbol{q}}\right)^{-\frac{1}{2}},$$

when $\boldsymbol{q} \to 1$. It follows that

$$(2.12)\quad \vartheta \sim \sqrt{\pi}\,\frac{S_{h,k}}{k}\left(\log\frac{1}{\boldsymbol{q}}\right)^{-\frac{1}{2}},$$

where

$$(2.121)\quad S_{h,k} = \sum_{j=1}^{k} e^{2j^2 h\pi i/k},$$

and

$$(2.13)\quad f(q) \sim \pi^{\frac{1}{2}s}\left(\frac{S_{h,k}}{k}\right)^s \left(\log\frac{1}{\boldsymbol{q}}\right)^{-\frac{1}{2}s},$$

it being understood that, when $S_{h,k} = 0$ (as is the case if, and only if, k is of the form $4m + 2$), this equation is to be understood as meaning

$$(2.131)\quad f(q) = o\left(\log\frac{1}{\boldsymbol{q}}\right)^{-\frac{1}{2}s}.$$

2.2. The principle of the method is to write down a power-series

$$f_{h,k}(q) = \sum c_{h,k,n}\, q^n, \tag{2.21}$$

which (a) is as simple and natural as possible and (b) behaves as much like $f(q)$ as possible when $q \to e^{2h\pi i/k}$; and to endeavor to approximate to the coefficients in $f(q)$ by means of the sums

$$\rho_s(n) = \sum_{h,k} c_{h,k,n}. \tag{2.22}$$

It is plain that, in forming these sums, we may ignore values of k of the form $4m+2$.

The appropriate auxiliary function (2.21) is

$$f_{h,k}(q) = \frac{\pi^{\frac{1}{2}s}}{\Gamma(\frac{1}{2}s)}\left(\frac{S_{h,k}}{k}\right)^s F_s(q), \tag{2.23}$$

where

$$F_s(q) = \sum_1^\infty n^{\frac{1}{2}s-1} q^n. \tag{2.231}$$

It is in fact well known that

$$F_s(x) - \Gamma(\tfrac{1}{2}s)\left(\log\frac{1}{x}\right)^{-\frac{1}{2}s}$$

is regular at $x = 1$.[7] We are thus led to take

$$c_{h,k,n} = \frac{\pi^{\frac{1}{2}s}\, n^{\frac{1}{2}s-1}}{\Gamma(\frac{1}{2}s)}\left(\frac{S_{h,k}}{k}\right)^s e^{-2nh\pi i/k} \tag{2.24}$$

and

$$\rho_s(n) = \sum_{h,k} c_{h,k,n} = \frac{\pi^{\frac{1}{2}s}\, n^{\frac{1}{2}s-1}}{\Gamma(\frac{1}{2}s)} \sum_1^\infty A_k, \tag{2.25}$$

where

$$A_1 = 1, \qquad A_k = k^{-s} \sum_h (S_{h,k})^s\, e^{-2nh\pi i/k}, \tag{2.251}$$

the summation extending over all positive values of h less than and prime to k. I call the series

$$\rho_s(n) = \frac{\pi^{\frac{1}{2}s}\, n^{\frac{1}{2}s-1}}{\Gamma(\frac{1}{2}s)} \sum A_k = \frac{\pi^{\frac{1}{2}s}\, n^{\frac{1}{2}s-1}}{\Gamma(\frac{1}{2}s)} S \tag{2.26}$$

the *singular series*. The process by which it has been constructed is of a purely formal character. It remains (1) to investigate more rigorously its bearing on the solution of our problem and (2) to find its sum.

3. Proof that the sum of the singular series, when $s = 8$ or $s = 5$, is the number of representations of n

3.1. *Proof that* $\rho_8(n) = r_8(n)$.

3.11. When s is 3, 4, 5, 6, 7, or 8 (but not 2 or any number greater

[7] See, for example, E. Lindelöf, *Le calcul des résidus*, p. 139.

than 8) the sum of the singular series gives exactly the number of representations of n. In this section I prove this when $s = 8$ and when $s = 5$. These cases are perfectly typical, but formally a little simpler than the others.

Suppose first that $s = 8$. Then

$$\Theta_8(q) = 1 + \sum f_{h,k}(q) = 1 + \frac{\pi^4}{6}\sum\left(\frac{S_{h,k}}{k}\right)^8 F_8(qe^{-2h\pi i/k}). \tag{3.111}$$

Now

$$S^8_{h,k} = \eta_k k^4,$$

where η_k is 1, 0, or 16 according as k is odd, oddly even, or evenly even. Also, if $x = e^{-y}$, we have

$$F_8(x) = \sum n^3 x^n = \sum n^3 e^{-ny} = \frac{d^2}{dy^2}(\tfrac{1}{4}\operatorname{cosech}^2 \tfrac{1}{2}y) = 6\sum \frac{1}{(y + 2n\pi i)^4},$$

where n runs through all integral values. Hence

$$F_8(qe^{-2h\pi i/k}) = \frac{6k^4}{\pi^4}\sum_n \frac{1}{\{2(nk + h) - k\tau\}^4},$$

and

$$\Theta_8(q) = 1 + \sum_{h,k,n} \frac{\eta_k}{\{2(nk + h) - k\tau\}^4}, \tag{3.112}$$

the summation extending over the values of h, k, and n already specified. If $k > 1$, $nk + h$ assumes all values prime to k; if $k = 1$, all values. Thus

$$\Theta_8(q) = 1 + \sum \frac{\eta_k}{(2h - k\tau)^4}; \tag{3.113}$$

where now $k = 1, 2, 3, \cdots$ and h assumes all values such that (h, k), the highest common factor of h and k, is unity. But this equation may be written

$$\begin{aligned}\Theta_8(q) &= 1 + \sum_{k=1,3,\ldots;\,(h,k)=1} \frac{1}{(2h - k\tau)^4} + \sum_{k=4,8,\ldots;\,(h,k)=1} \frac{16}{(2h - k\tau)^4} \\ &= 1 + \sum_{k=1,3,\ldots;\,(h,k)=1} \frac{1}{(2h - k\tau)^4} + \sum_{k=2,4,\ldots;\,(h,k)=1} \frac{1}{(h - k\tau)^4} \\ &= 1 + \sum \frac{1}{(h - k\tau)^4},\end{aligned} \tag{3.114}$$

where now $k = 1, 2, 3, \cdots$ and h assumes all values *of opposite parity to and prime to* k.

Multiplying both sides of (3.114) by

$$\frac{\pi^4}{96} = 1 + \frac{1}{3^4} + \frac{1}{5^4} + \cdots,$$

we obtain

$$\frac{\pi^4}{96}\Theta_8(q) = \frac{\pi^4}{96} + \sum \frac{1}{(h - k\tau)^4} \tag{3.1151}$$

or

$$\frac{\pi^4}{96}\Theta_3(q) = \sum \frac{1}{(h-k\tau)^4}. \tag{3.1152}$$

In (3.1151) $k = 1, 2, 3, \cdots$, and in (3.1152) $k = 0, 1, 2, \cdots$; in each equation h assumes *all* values of opposite parity to k.

3.12. We now write

$$\frac{\pi^4}{96}\Theta_3(q) = \chi(\tau), \tag{3.121}$$

and consider the effect on $\chi(\tau)$ of the modular substitutions

$$\tau' = \tau \pm 2, \tag{3.1221}$$

$$\tau' = -1/\tau. \tag{3.1222}$$

It is obvious in the first place, from (3.1151) or (3.1152), that

$$\chi(\tau \pm 2) = \chi(\tau). \tag{3.123}$$

Again, we may write (3.1151) in the form

$$\chi(\tau) = \frac{\pi^4}{96} + \frac{1}{2}\sum_k{}' \sum_h \frac{1}{(h-k\tau)^4},$$

where h and k assume all values of opposite parity except that (as is indicated by the dash) the value $k = 0$ is omitted. Thus

$$\chi\left(-\frac{1}{\tau}\right) = \frac{\pi^4}{96} + \frac{1}{2}\tau^4 \sum_k{}' \sum_h \frac{1}{(h\tau+k)^4} = \frac{\pi^4\tau^4}{96} + \frac{1}{2}\tau^4 \sum_k \sum_h{}' \frac{1}{(h\tau+k)^4}.$$

Changing h and k into $-k$ and h, we obtain

$$\chi(-1/\tau) = \tau^4 \chi(\tau). \tag{3.124}$$

Now $\{\vartheta_3(0,\tau)\}^8 = \vartheta^8 = \psi(\tau)$ satisfies the equations

$$\psi(\tau \pm 2) = \psi(\tau), \qquad \psi(-1/\tau) = \tau^4 \psi(\tau);$$

and so it follows, from (3.123) and (3.124), that the function

$$\eta(\tau) = \chi(\tau)/\vartheta^8 = \chi(\tau)/\psi(\tau) \tag{3.125}$$

is invariant for the substitutions (3.122), and therefore for the modular subgroup which they generate, the group called by Klein-Fricke and Mordell Γ_3.

3.13. The next step in the proof is to show that $\eta(\tau)$ is bounded throughout the "fundamental polygon" G_3 associated with the group Γ_3. This region is defined by

$$\tau = x + iy, \qquad |\tau| \geqq 1, \qquad -1 \leqq x \leqq 1,$$

and has only the points $\tau = \pm 1$ in common with the real axis. It is therefore sufficient to show that $\eta(\tau)$ is bounded when τ approaches one or other

of these points, say $\tau = 1$. For this purpose, following Mordell, I consider the effect of the substitution

$$\tau = 1 - \frac{1}{T}.$$

If we write $T = X + iY$, and suppose that $\tau \to 1$ from inside G_3, then $Y \to \infty$ and $|Q| = |e^{\pi i T}|$ is small. And

$$\chi\left(1 - \frac{1}{T}\right) = \frac{\pi^4}{96} + \frac{1}{2} T^4 \sum_k{}' \sum_h \frac{1}{\{k + (h-k)T\}^4} \tag{3.131}$$

$$= \frac{\pi^4}{96} + \frac{1}{2} T^4 \sum_k{}' \sum_h \frac{1}{(k + hT)^4}$$

$$= \frac{1}{2} T^4 \sum_k \sum_h \frac{1}{(k + hT)^4},$$

where now k assumes all integral values and h all odd values.

Write $hT = a$, $e^{\pi i a} = \zeta$, and sum with respect to k. We have

$$\sum_k \frac{1}{(k+a)^4} = -\frac{2}{3}\pi^4 \operatorname{cosec}^2 a\pi + \pi^4 \operatorname{cosec}^4 a\pi;$$

and this function, when expanded in powers of ζ, begins with the term

$$\tfrac{8}{3}\pi^4 \zeta^2 = \tfrac{8}{3}\pi^4 Q^{2h}.$$

Hence, when $T = X + iY$ and Y is large, we have

$$\chi\left(1 - \frac{1}{T}\right) = \frac{8}{3}\pi^4 T^4 Q^2 + \cdots,$$

$$\Theta_3(q) = 256 T^4 Q^2 + \cdots. \tag{3.132}$$

But we have also

$$\left\{\vartheta_3\left(0, 1 - \frac{1}{T}\right)\right\}^8 = T^4\{\vartheta_2(0, T)\}^8 = 256 T^4 Q^2 + \cdots, \tag{3.133}$$

and so

$$\eta(\tau) = \eta\left(1 - \frac{1}{T}\right) \to 1. \tag{3.134}$$

Thus $\eta(\tau)$ is an invariant of Γ_3 and bounded throughout G_3; and is therefore necessarily a constant, which is plainly unity.

It follows that

$$\vartheta^8 = \Theta_3(q) \tag{3.135}$$

and so that

$$\rho_8(n) = r_8(n). \tag{3.136}$$

3.2. *Proof that* $\rho_5(n) = r_5(n)$.

3.21. When $s = 5$ the proof proceeds on the same lines, but is not quite so simple. We shall require certain well-known identities which I state as lemmas.

LEMMA 3.211.[8] *If h and k are positive integers of opposite parity, then*

$$\sum_{j=1}^{k} e^{j^2 h\pi i/k} = \sqrt{i}\sqrt{\frac{k}{h}}\sum_{j=1}^{h} e^{-j^2 k\pi i/h}; \tag{3.2111}$$

and if h and k are positive integers, and h odd, then

$$\sum_{j=1}^{k} (-1)^j e^{j^2 h\pi i/k} = \sqrt{i}\sqrt{\frac{k}{h}}\sum_{j=1}^{h} e^{-(j-\frac{1}{2})^2 k\pi i/h}. \tag{3.2112}$$

Here $\sqrt{i} = e^{\frac{1}{4}\pi i}$.

LEMMA 3.212. *Suppose that $0 < \nu < 1$, and that σ and the real part of t are positive. Then*

$$\frac{(2\pi)^\sigma}{\Gamma(\sigma)}\sum_{0}^{\infty}(m+\nu)^{\sigma-1}e^{-2\pi t(m+\nu)} = \sum_{-\infty}^{\infty}\frac{e^{2n\nu\pi i}}{(t+ni)^\sigma}, \tag{3.2121}$$

where

$$(t+ni)^\sigma = exp\{\sigma \log (t+ni)\} = exp(\sigma \log |t+ni| + \sigma\phi i)$$

and $-\frac{1}{2}\pi < \phi < \frac{1}{2}\pi$. The formula still holds for $\nu = 0$, if $\sigma > 1$.

This result is due to Lipschitz.[9] We shall require two special cases.

(i) Suppose that $\sigma = \frac{1}{2}s > 1$, $\nu = 0$, $t = -\frac{1}{2}i\tau$, $x = e^{\pi i\tau}$; so that $\Im(\tau) > 0$ and $|x| < 1$. Then we obtain

$$\frac{\pi^{\frac{1}{2}s}}{\Gamma(\frac{1}{2}s)}F_s(x) = \sum_{-\infty}^{\infty}\frac{1}{\{(2n-\tau)i\}^{\frac{1}{2}s}}, \tag{3.2122}$$

where

$$\{(2n-\tau)i\}^{\frac{1}{2}s} = exp\{\tfrac{1}{2}s \log |(2n-\tau)i| + \tfrac{1}{2}s\phi i\} \tag{3.21221}$$

and $-\frac{1}{2}\pi < \phi < \frac{1}{2}\pi$.

(ii) Suppose that $\nu = \frac{1}{2}(1+\theta)$, $\theta = \lambda/K$, where K and λ are integers and $-K < \lambda < K$, and $t = -K\tau i$. Then

$$\sum_{-\infty}^{\infty}\frac{(-1)^n e^{n\theta\pi i}}{\{(n-K\tau)i\}^\sigma} = x^{K+\lambda}P(x), \tag{3.2123}$$

where $P(x)$ is an ascending power series in x.

3.22. Supposing now that $s = 5$, we have

$$\Theta_5(q) = 1 + \frac{4\pi^2}{3}\sum\left(\frac{S_{h,k}}{k}\right)^5 F_5(qe^{-2h\pi i/k}), \tag{3.221}$$

[8] See, for example, G. Landsberg, *Zur Theorie der Gaussschen Summen und der linearen Transformation der Thetafunktionen*, Journal für Mathematik,, vol. 111 (1893), pp. 234–253. Both formulas are included in formula (17*b*), p. 243, of Landsberg's memoir. The first is also proved by Lindelöf, loc. cit., pp. 73–75.

[9] R. Lipschitz, *Untersuchung der Eigenschaften einer Gattung von unendlichen Reihen*, Journal für Mathematik, vol. 105 (1889), pp. 127–156.

and

$$S_{h,k}^4 = \eta_k k^2,$$

where now η_k is 1, 0, or -4 according as k is odd, oddly even, or evenly even. Thus

$$\Theta_5(q) = 1 + \frac{4\pi^2}{3} \sum \frac{\eta_k}{k^3} S_{h,k} F_5(qe^{-2h\pi i/k}). \tag{3.222}$$

Substituting from (3.2122), we obtain

$$\Theta_5(q) = 1 + \sum \frac{\eta_k}{\sqrt{k}} \frac{S_{h,k}}{[\{2(nk+h) - k\tau\}i]^{5/2}}, \tag{3.223}$$

or

$$\Theta_5(q) = 1 + \sum \frac{\eta_k}{\sqrt{k}} \frac{S_{h,k}}{\{(2h - k\tau)i\}^{5/2}}, \tag{3.224}$$

the ranges of summation in these equations being the same as in (3.112) and (3.113) respectively. The last equation can be expressed in a more convenient form by introducing the sum

$$T_{h,k} = \sum_0^{k-1} e^{j2h\pi i/k}. \tag{3.225}$$

In fact (3 224) may be written

$$\Theta_5(q) = 1 + \sum_{k=1,3,\cdots;\,(h,k)=1} \frac{1}{\sqrt{k}} \frac{S_{h,k}}{\{(2h - k\tau)i\}^{5/2}} - \sum_{k=4,8,\cdots;\,(h,k)=1} \frac{4}{\sqrt{k}} \frac{S_{h,k}}{\{(2h - k\tau)i\}^{5/2}}.$$

In the first sum k is odd; $2h = H$ is even and prime to k; and $S_{h,k} = T_{H,k}$. In the second $k = 2K$, where K runs through all even values; h is odd and prime to K; and $S_{h,k} = 2T_{h,K}$. Effecting these substitutions, and then replacing H (or K) by h (or k), we obtain

$$\Theta_5(q) = 1 + \sum \frac{(-1)^h}{\sqrt{k}} \frac{T_{h,k}}{\{(h - k\tau)i\}^{5/2}},$$

where now $k = 1, 2, 3, \cdots$ and h assumes all values of opposite parity and prime to k.[10]

Multiplying by

$$\frac{\pi^2}{8} = \frac{1}{1^2} + \frac{1}{3^2} + \frac{1}{5^2} + \cdots,$$

and observing that, if λ is odd,

$$(-1)^{\lambda h} = (-1)^h, \qquad \sqrt{\lambda k}\{(\lambda h - \lambda k\tau)i\}^{5/2} = \lambda^3 \sqrt{k}\{(h - k\tau)i\}^{5/2},$$

$$T_{\lambda h, \lambda k} = \lambda T_{h,k},$$

[10] This is equation (8) of my former paper.

we obtain

$$\frac{\pi^2}{8}\Theta_5(q) = \frac{\pi^2}{8} + \sum \frac{(-1)^h}{\sqrt{k}} \frac{T_{h,k}}{\{(h-k\tau)i\}^{5/2}}, \tag{3.226}$$

where h now assumes all values of opposite parity to k.[11]

3.23. The discussion now follows the lines of 3.12 and 3.13. We write

$$\frac{\pi^2}{8}\Theta_5(q) = \chi(\tau), \tag{3.231}$$

and it is obvious at once that

$$\chi(\tau \pm 2) = \chi(\tau). \tag{3.232}$$

The discussion of the transformation $\tau' = -1/\tau$ requires a little more care, owing to the presence of many-valued functions under the sign of summation. It is convenient to begin by including negative values of k.

We write generally

$$z^s = exp\{s \log |z| + i \operatorname{am} z\}$$

where the particular value of $\operatorname{am} z$ to be selected has to be fixed by special convention. Thus in $\{(h-k\tau)i\}^{5/2}$, where $k>0$, $\operatorname{am}\{(h-k\tau)i\}$ lies (as has already been explained) between $-\frac{1}{2}\pi$ and $\frac{1}{2}\pi$. We now agree that, if k is still positive, $\operatorname{am}(-k) = \pi$, so that $\sqrt{-k} = i\sqrt{k}$, while $\operatorname{am}\{(-h+k\tau)i\}$ lies between $-\frac{3}{2}\pi$ and $-\frac{1}{2}\pi$. It will easily be verified that

$$\sqrt{-k}\{(-h+k\tau)i\}^{5/2} = \sqrt{k}\{(h-k\tau)i\}^{5/2}. \tag{3.233}$$

Further, we write by definition

$$T_{-h,-k} = T_{h,k}. \tag{3.234}$$

We know from (3.2111) that, when h and k are both positive,

$$T_{h,k} = \sqrt{i}\sqrt{\frac{k}{h}}\, T_{-k,h};$$

and it is easy to verify that, with our conventions, we have generally

$$T_{h,k} = \epsilon\sqrt{i}\sqrt{\frac{k}{h}}\, T_{-k,h}, \tag{3.235}$$

where $\epsilon = 1$ unless $h>0$, $k<0$, in which case $\epsilon = -1$.

3.24. We have, from (3.226), (3.233), and (3.234),

$$\begin{aligned}\chi(\tau) &= \frac{\pi^2}{8} + \frac{1}{2}\sum{}' \sum_h \frac{(-1)^h}{\sqrt{k}} \frac{T_{h,k}}{\{(h-k\tau)i\}^{5/2}} \\ &= \frac{\pi^2}{8} + \frac{\pi^2}{8}(-\tau i)^{-5/2} + \frac{1}{2}\sum{}' \frac{(-1)^h}{\sqrt{k}} \frac{T_{h,k}}{\{(h-k\tau)i\}^{5/2}},\end{aligned} \tag{3.241}$$

[11] This equation takes the place of (9) of my former paper, which is not printed correctly. The first term on the right is omitted, and $k=0$ is included wrongly under the sign of summation.

where now h and k are any integers other than zero and of opposite parity. Writing $-1/\tau$ for τ in (3.241), using (3.235), and then replacing h and k by K and $-H$, we obtain

$$\chi\left(-\frac{1}{\tau}\right)=\frac{\pi^2}{8}+\frac{\pi^2}{8}\left(\frac{i}{\tau}\right)^{-5/2}+\frac{1}{2}\sqrt{i}\,{\sum}'\frac{(-1)^K\epsilon}{\sqrt{K}}\frac{T_{H,K}}{\{(K-H/\tau)\,i\}^{5/2}}, \tag{3.242}$$

where ϵ is 1 unless H and K are both positive, and then -1, and

$$\operatorname{am}\{(K-H/\tau)\,i\}$$

lies between $-\frac{1}{2}\pi$ and $\frac{1}{2}\pi$ if $H<0$, between $-\frac{3}{2}\pi$ and $-\frac{1}{2}\pi$ if $H>0$.

It may be verified without difficulty that

$$\left\{\left(K-\frac{H}{\tau}\right)i\right\}^{5/2}=\epsilon\left(-\frac{1}{\tau}\right)^{5/2}\{(H-K\tau)\,i\}^{5/2}, \tag{3.243}$$

where $0<\operatorname{am}(-1/\tau)<\pi$ and the value of $\operatorname{am}\{(H-K\tau)\,i\}$ is fixed in accordance with our previous conventions. Consider, for example, the case $H>0$, $K<0$. In this case

$$-\tfrac{3}{2}\pi<\alpha=\operatorname{am}\{(K-H/\tau)\,i\}<-\tfrac{1}{2}\pi,$$

$$0<\beta=\operatorname{am}(-1/\tau)<\pi,$$

and

$$-\tfrac{1}{2}\pi<\gamma=\operatorname{am}\{(H-K\tau)\,i\}<\tfrac{1}{2}\pi.$$

Thus $\beta+\gamma$ lies between $-\frac{1}{2}\pi$ and $\frac{3}{2}\pi$, and, as α differs from $\beta+\gamma$ by a multiple of 2π, we must have $\alpha=\beta+\gamma-2\pi$ and

$$\left\{\left(K-\frac{H}{\tau}\right)i\right\}^{5/2}=e^{-5\pi i}\left(-\frac{1}{\tau}\right)^{5/2}\{(H-K\tau)\,i\}^{5/2}$$

$$=-\left(-\frac{1}{\tau}\right)^{5/2}\{(H-K\tau)\,i\}^{5/2},$$

in agreement with (3.243). The other possible cases may be treated similarly.

Thus

$$\chi\left(-\frac{1}{\tau}\right)=\frac{\pi^2}{8}+\frac{\pi^2}{8}\left(\frac{i}{\tau}\right)^{-5/2} \tag{3.244}$$

$$+\tfrac{1}{2}\sqrt{i}\left(-\frac{1}{\tau}\right)^{-5/2}{\sum}'\frac{(-1)^K}{\sqrt{K}}\frac{T_{H,K}}{\{(H-K\tau)\,i\}^{5/2}},$$

where $-\frac{1}{2}\pi<\operatorname{am}(i/\tau)<\frac{1}{2}\pi$, $0<\operatorname{am}(-1/\tau)<\pi$. And this equation

leads to

$$(3.245)\quad \chi\left(-\frac{1}{\tau}\right)=\frac{\pi^2}{8}\left(\frac{i}{\tau}\right)^{-5/2}+\tfrac{1}{2}\sqrt{i}\left(-\frac{1}{\tau}\right)^{-5/2}\sum_K{}'\sum_H\frac{(-1)^K}{\sqrt{K}}\frac{T_{H,K}}{\{(H-K\tau)i\}^{5/2}}$$

$$=\left(\frac{\tau}{i}\right)^{5/2}\left[\frac{\pi^2}{8}+\frac{1}{2}\sum_K{}'\sum_H\frac{(-1)^H}{\sqrt{K}}\frac{T_{H,K}}{\{(H-K\tau)i\}^{5/2}}\right]$$

$$=\left(\frac{\tau}{i}\right)^{5/2}\chi(\tau).^{12}$$

This is the same functional equation as is satisfied by ϑ^5. Hence

$$(3.246)\qquad \eta(\tau)=\chi(\tau)/\vartheta^5$$

is an invariant for each of the substitutions (3.122), and so for Γ_3.

3.25. It remains to verify that $\eta(\tau)$ is bounded in G_3. As in 3.13, it is only necessary to consider the neighborhood of $\tau=1$. Putting $\tau=1-1/T$, as in 3.13, in (3.241), and then writing $h=H-K$, $k=H$, we obtain

$$(3.251)\quad \frac{\pi^2}{8}\chi\left(1-\frac{1}{T}\right)=\frac{\pi^2}{8}+\frac{1}{2}\sum_H{}'\sum_K\frac{(-1)^H}{\sqrt{H}}\frac{T_{H-K,H}}{\{(-K+H/\tau)i\}^{5/2}},$$

where H assumes all values save 0 and K all odd values, and

$$\operatorname{am}\{(-K+H/\tau)i\}$$

lies between $-\frac{1}{2}\pi$ and $\frac{1}{2}\pi$, or between $-\frac{3}{2}\pi$ and $-\frac{1}{2}\pi$, according as $H>0$ or $H<0$.

Now

$$(3.252)\qquad T_{H-K,H}=\sum_0^{H-1}(-1)^j e^{-j^2K\pi i/H}=U_{K,H},$$

say. It follows from (3.2112) that, if h and k are both positive, and h is odd, and

$$(3.253)\qquad U_{h,k}=\sum_0^{k-1}(-1)^j e^{-j^2h\pi i/k},\qquad W_{h,k}=\sum_1^{k}e^{(j-\frac{1}{2})^2h\pi i/k},$$

then

$$U_{h,k}=\sqrt{-i}\sqrt{\frac{k}{h}}W_{k,h};$$

and it is easily verified that, if we adopt the same conventions as in 3.23 concerning the meanings of $\sqrt{-h}$, $\sqrt{-k}$, $U_{-h,-k}$, and $W_{-h,-k}$, we have

[12] We first restore the terms for which $H=0$, and then observe that

$$(-1)^H=-(-1)^K.$$

We have to verify that, with our conventions,

$$-\sqrt{i}\,(-1/\tau)^{-\frac{1}{2}}(-\tau i)^{-\frac{1}{2}}=1$$

and

$$-\sqrt{i}\,(-1/\tau)^{-\frac{1}{2}}=-(i/\tau)^{-\frac{1}{2}};$$

these verifications present no difficulty.

generally

$$U_{h,\,k} = \epsilon\sqrt{-i}\sqrt{\frac{k}{h}}\,W_{k,\,h}, \tag{3.254}$$

where $\epsilon = 1$ unless $h < 0$, $k > 0$, in which case $\epsilon = -1$.

Using this equation in (3.251), we obtain

$$\frac{\pi^2}{8}\chi\left(1-\frac{1}{T}\right) = \frac{\pi^2}{8} - \tfrac{1}{2}\sqrt{-i}\sum_H{}'\sum_K \frac{(-1)^H\epsilon}{\sqrt{K}}\frac{W_{H,K}}{\{(-K+H/T)i\}^{5/2}} \tag{3.255}$$

$$= -\tfrac{1}{2}\sqrt{-i}\sum_H\sum_K \frac{(-1)^H\epsilon}{\sqrt{K}}\frac{W_{H,K}}{\{(-K+H/T)i\}^{5/2}}.$$

It is now easy to verify, by arguments similar to those of 3.24, that

$$\{(-K+H/T)i\}^{5/2} = \epsilon T^{-5/2}\{(H-KT)i\}^{5/2},$$

where $\operatorname{am} T$ lies between 0 and π, while $\operatorname{am}\{(H-KT)i\}$ obeys our previous conventions. We thus obtain

$$\frac{\pi^2}{8}\chi\left(1-\frac{1}{T}\right) = -\tfrac{1}{2}\sqrt{-i}\,T^{5/2}\sum_H\sum_K \frac{(-1)^H}{\sqrt{K}}\frac{W_{H,K}}{\{(H-KT)i\}^{5/2}},$$

the summation being now limited only by the fact that K is odd. In virtue of (3.233), this equation may be written

$$\frac{\pi^2}{8}\chi\left(1-\frac{1}{T}\right) = -\sqrt{-i}\,T^{5/2}\sum_{K=1,\,3,\,5,\,\ldots}\frac{1}{\sqrt{K}}\sum_H\frac{(-1)^H W_{H,K}}{\{(H-KT)i\}^{5/2}}. \tag{3.256}$$

3.26. The series in (3.256) may be expressed in the form

$$\sum_{K=1,\,3,\,5,\,\ldots}\frac{1}{\sqrt{K}}\sum_{j=1}^{K}\sum_H\frac{(-1)^H e^{-(j-\frac{1}{2})^2H\pi i/K}}{\{(H-KT)i\}^{5/2}}.$$

Suppose that

$$(j-\tfrac{1}{2})^2 \equiv \lambda_j \pmod{2K} \qquad (-K < \lambda_i < K),$$

and $\theta_j = \lambda_j/K$. Then

$$\sum_H\frac{(-1)^H e^{-(j-\frac{1}{2})^2H\pi i/K}}{\{(H-KT)i\}^{5/2}} = \sum_H\frac{(-1)^H e^{-H\theta_j\pi i}}{\{(H-KT)i\}^{5/2}}$$

may, by (3.2123), be expanded as a power series in $Q = e^{\pi i T}$, in which the lowest power of Q is

$$Q^{K+\lambda_j}.$$

The smallest possible values of $K+\lambda_j$ are $\frac{1}{4}, \frac{5}{4}, \cdots$; and $K+\lambda_j = \frac{1}{4}$ involves

$$j^2 - j + \tfrac{1}{4} \equiv -K + \tfrac{1}{4} \pmod{2K}$$

or

$$j^2 - j = (2p-1)K,$$

where p is an integer, i.e. an equation whose left-hand side is even and whose right-hand side is odd. Thus $K + \lambda_j \geqq \frac{5}{4}$; and the left-hand side of (3.256) is the product of $T^{5/2} Q^{5/4}$ by an ascending power-series in Q. But

$$\vartheta_3\left(0, 1 - \frac{1}{T}\right) = \frac{\sqrt{T}}{\sqrt{i}} \vartheta_2(0, T)$$

is the product of $T^{1/2} Q^{1/4}$ by a power-series in Q. It now follows, just as in 3.13, that $\eta(\tau)$ is bounded, and so is a constant, which is plainly unity.

We have thus established the identity of Θ_s and ϑ^s, and so of $\rho_s(n)$ and $r_s(n)$, when $s = 8$ and $s = 5$. The same method may be used for any value of s from 5 to 8 inclusive.[13] In order to complete the solution of our problem, we have to sum the singular series (2.26).

4. General rules for the summation of the singular series

4.1. *The value of* $S_{h,k}$.

The known results concerning the value of the Gaussian sum $S_{h,k}$ are as follows.[14] We assume that $(h, k) = 1$.

If k and k' are prime to one another

$$S_{h,kk'} = S_{hk',k} S_{hk,k'}. \tag{4.11}$$

We need therefore consider only the cases in which $k = 2^\lambda$ or $K = p^\lambda$, p being an odd prime.

If $k = 2$

$$S_{h,k} = 0. \tag{4.121}$$

If $k = 2^\lambda = 2^{2\mu+1}$, and $\mu > 0$,

$$S_{h,k} = 2^{\mu+1} e^{\frac{1}{4}h\pi i}. \tag{4.122}$$

If $k = 2^\lambda = 2^{2\mu}$, and $\mu > 0$,

$$S_{h,k} = 2^\mu (1 + i^h) = 2^{\mu+1} \cos \tfrac{1}{4}h\pi \, e^{\frac{1}{4}h\pi i}. \tag{4.123}$$

If $k = p$,

$$S_{h,k} = \left(\frac{h}{p}\right) i^{\frac{1}{4}(p-1)^2} \sqrt{p}, \tag{4.131}$$

where (h/p) is the well known symbol of Legendre and Jacobi.

[13] When $s = 2$ or $s > 8$, the conclusion is false. The cases $s = 3$ and $s = 4$ are exceptional. The conclusion is true, but new difficulties arise in the proof because the series used are not all absolutely convergent. These difficulties are easily surmounted when $s = 4$, but are more serious when $s = 3$.

[14] For proofs of these assertions see the chapter on Gauss's sums in the second volume of Bachmann's *Zahlentheorie*. A less complete account is given in Dirichlet-Dedekind, *Vorlesungen über Zahlentheorie*, ed. 4, 1894, pp. 287 et seq.

If $k = p^\lambda = p^{2\mu+1}$, and $\mu > 0$,

$$S_{h,\,k} = p^\mu\, S_{h,\,p}. \tag{4.132}$$

Finally, if $k = p^\lambda = p^{2\mu}$, and $\mu > 0$,

$$S_{h,\,k} = p^\mu. \tag{4.133}$$

These formulas enable us to write down the value of $S_{h,\,k}$ for all co-prime pairs of values of h and k. .

The multiplication rule for A_k.

4.2. The first step is to prove that

$$A_{kk'} = A_k\, A_{k'} \tag{4.21}$$

whenever k and k' are prime to one another.

In the formula which defines A_k, viz.

$$k^s\, A_k = \sum (S_{h,\,k})^s\, e^{-2nh\pi i/k},$$

h assumes all positive values less than and prime to k. Let us call this set of values, or any set congruent to this set to modulus k, a *k-set.* It is easy to see that if h runs through a k-set, and h' through a k'-set, then

$$\boldsymbol{h} = hk' + h'\,k$$

runs through a kk'-set. For the number of values of $\boldsymbol{h}$ is

$$\phi(k)\,\phi(k') = \phi(kk'),$$

and it is obvious that all are prime to kk' and incongruent to modulus kk'.

Thus

$$\begin{aligned}(kk')^s\, A_{kk'} &= \sum_h (S_{h,\,kk'})^s\, e^{-2nh\pi i/kk'} \\ &= \sum_h (S_{hk',\,k})^s\, (S_{hk,\,k'})^s\, e^{-2nh\pi i/kk'},\end{aligned}$$

by (4.11). But

$$S_{hk',\,k} = \sum_{j=1}^{k} e^{2j^2hk'\pi i/k} = \sum_{j=1}^{k} e^{2(jk')^2h\pi i/k} = S_{h,\,k},$$

since jk' runs through a complete system of residues to modulus k; and similarly $S_{hk,\,k'} = S_{h',\,k'}$. Thus

$$\begin{aligned}(kk')^s\, A_{kk'} &= \sum_{h,\,h'} (S_{h,\,k})^s\, (S_{h',\,k'})^s\, e^{-2nh\pi i/kk'} \\ &= \sum_h (S_{h,\,k})^s\, e^{-2nh\pi i/k} \sum_{h'} (S_{h',\,k'})^s\, e^{-2nh'\pi i/k'} \\ &= (kk')^s\, A_k\, A_{k'};\end{aligned}$$

which proves (4.21).

It follows[15] that

$$S = A_1 + A_2 + A_3 + \cdots = 1 + A_2 + A_3 + \cdots = \prod \chi_p \tag{4.22}$$

where

$$\chi_p = 1 + A_p + A_{p^2} + A_{p^3} + \cdots \tag{4.23}$$

and p runs through all prime values.

Calculation of A_{2^λ}.

4.3. Suppose first that $p = 2$. Then the value of A_{2^λ} is given by the following system of rules.

4.31. $$A_2 = 0.$$

4.32. *If λ is odd and greater than 1,*

$$A_{2^\lambda} = 0, \tag{4.321}$$

unless

$$n \equiv 0 \pmod{2^{\lambda-3}} \tag{4.3221}$$

and

$$\nu - s = 2^{-(\lambda-3)}\, n - s \equiv 0 \pmod 4, \tag{4.3222}$$

in which case

$$A_{2^\lambda} = 2^{-(\frac{1}{2}s-1)(\lambda-1)}\, e^{-\frac{1}{4}(\nu-s)\pi i}. \tag{4.3223}$$

If $\lambda = 3$, (4.3221) is satisfied automatically.

Let $\lambda = 2\mu + 1\,(\mu > 0)$. Then

$$S_{h,\,2^\lambda} = 2^{\mu+1}\, e^{\frac{1}{4}h\pi i} = 2^{\mu-1}\, S_{h,\,8},$$

by (4.122). We write

$$h = 8z + h' \qquad (z = 0, 1, \cdots, 2^{\lambda-3} - 1;\ h' = 1, 3, 5, 7).$$

Then

$$A_{2^\lambda} = 2^{-s(\mu+2)} \sum_h (S_{h,\,8})^s\, e^{-2nh\pi i/2^{2\mu+1}}$$

$$= 2^{-s(\mu+2)} \sum_{h'} (S_{h',\,8})^s\, e^{-2nh'\pi i/2^{2\mu+1}} \sum_z e^{-2nz\pi i/2^{2\mu-2}},$$

which vanishes (in virtue of the summation with respect to z) unless $n \equiv 0 \pmod{2^{2\mu-2}}$.

If $n = 2^{2\mu-2}\,\nu$, we have

$$A_{2^\lambda} = 2^{2\mu-2-s(\mu+2)} \sum_{h'} (S_{h',\,8})^s\, e^{-\frac{1}{4}\nu h'\pi i}.$$

The sum with respect to h' is

$$2^{2s}\{e^{-\frac{1}{4}(\nu-s)\pi i} + e^{-\frac{3}{4}(\nu-s)\pi i} + e^{-\frac{5}{4}(\nu-s)\pi i} + e^{-\frac{7}{4}(\nu-s)\pi i}\},$$

[15] We assume that the series and product are absolutely convergent. This is obviously the case if $s > 4$, as $S_{h,k} = O(\sqrt{k})$, $A_k = O(k^{1-\frac{1}{2}s})$, and $1 - \frac{1}{2}s < -1$.

which is 0 or $2^{2s+2}\,e^{-\frac{1}{2}(\nu-s)\pi i}$ according as (4.3222) is not or is satisfied. This completes the proof of 4.32.

4.33. *If λ is even and greater than 1,*

$$A_{2^\lambda} = 0 \tag{4.331}$$

unless

$$n \equiv 0 \pmod{2^{\lambda-2}}, \tag{4.3321}$$

in which case

$$A_{2^\lambda} = 2^{-(\frac{1}{2}s-1)(\lambda-1)} \cos\left(\tfrac{1}{2}\nu\pi - \tfrac{1}{4}s\pi\right), \tag{4.3322}$$

where $n = 2^{\lambda-2}\nu$. If $\lambda = 2$, the last formula holds in any case.

If $\lambda = 2\mu$, we have

$$S_{h,\,2^\lambda} = 2^{\mu+1} \cos\tfrac{1}{4}h\pi\, e^{\frac{1}{2}h\pi i} = 2^{\mu-1} S_{h,\,4},$$

by (4.123). We write

$$h = 4z + h' \qquad (z = 0, 1, \cdots, 2^{\lambda-2} - 1;\ h' = 1, 3).$$

Then

$$A_{2^\lambda} = 2^{-s(\mu+1)} \sum_h (S_{h,\,4})^s\, e^{-2nh\pi i/2^{2\mu}}$$

$$= 2^{-s(\mu+1)} \sum_{h'} (S_{h',\,4})^s\, e^{-2nh'\pi i/2^{2\mu}} \sum_z e^{-2nz\mu i/2^{2\mu-2}},$$

which vanishes (in virtue of the summation with respect to z) unless $n \equiv 0 \pmod{2^{2\mu-2}}$. But if $n = 2^{2\mu-2}\nu$, we have

$$A_{2^\lambda} = 2^{2\mu-2-s(\mu+1)} \sum_{h'} (S_{h',\,4})^s\, e^{-\frac{1}{2}\nu h'\pi i},$$

and the sum here is

$$2^{2s}\{(\cos\tfrac{1}{4}\pi)^s\, e^{\frac{1}{2}s\pi i - \frac{1}{2}\nu\pi i} + (\cos\tfrac{3}{4}\pi)^s\, e^{\frac{1}{2}s\pi i - \frac{3}{2}\nu\pi i}\} = 2^{\frac{1}{2}s+1} \cos\left(\tfrac{1}{2}\nu\pi - \tfrac{1}{4}s\pi\right).$$

This completes the proof of 4.33. If $\lambda = 2$, z disappears from the argument and h and h' are identical.

Calculation of A_{p^λ} when p is odd.

4.4. The corresponding results when p is odd are as follows.

4.41. *If $n \not\equiv 0 \pmod{p}$ then*

$$A_p = -\,p^{-\frac{1}{2}s} \qquad (s \equiv 0), \tag{4.411}$$

$$A_p = \left(\frac{n}{p}\right) p^{-\frac{1}{2}(s-1)} \qquad (s \equiv 1), \tag{4.412}$$

$$A_p = -\,(-1)^{\frac{1}{4}(p-1)}\, p^{-\frac{1}{2}s} \qquad (s \equiv 2), \tag{4.413}$$

$$A_p = (-1)^{\frac{1}{4}(p-1)} \left(\frac{n}{p}\right) p^{-\frac{1}{2}(s-1)} \qquad (s \equiv 3); \tag{4.414}$$

the congruences for s referring to modulus 4. But if $n \equiv 0 \pmod{p}$ then

$$A_p = (p-1)\, p^{-\frac{1}{2}s} \qquad (s \equiv 0), \tag{4.415}$$

$$A_p = 0 \qquad (s \equiv 1, 3), \tag{4.416}$$

$$A_p = (-1)^{\frac{1}{2}(p-1)}(p-1)p^{-\frac{1}{2}s} \qquad (s \equiv 2). \tag{4.417}$$

We have

$$A_p = p^{-s}\sum_h (S_{h,p})^s e^{-2nh\pi i/p} = i^{\frac{1}{2}s(p-1)^2} p^{-\frac{1}{2}s}\sum_h \left(\frac{h}{p}\right)^s e^{-2nh\pi i/p}.$$

If s is even, this is

$$i^{\frac{1}{2}s(p-1)^2} p^{-\frac{1}{2}s}\sum_h e^{-2nh\pi i/p},$$

and the sum is equal to -1 or to $p-1$ according as n is not or is a multiple of p. This leads at once to the results stated for even values of s.

If on the other hand s is odd, we have

$$A_p = i^{\frac{1}{2}s(p-1)^2} p^{-\frac{1}{2}s}\sum_h \left(\frac{h}{p}\right) e^{-2nh\pi i/p},$$

which is equal to 0 if n is a multiple of p, and to

$$i^{\frac{1}{2}(s-1)(p-1)^2}\left(\frac{n}{p}\right) p^{-\frac{1}{2}(s-1)}$$

otherwise. We thus obtain the results stated for odd values of s.

4.42. *If λ is odd and greater than 1, then*

$$A_{p^\lambda} = 0 \tag{4.421}$$

if $n \not\equiv 0 \pmod{p^{\lambda-1}}$;

$$A_{p^\lambda} = p^{-(\frac{1}{2}s-1)(\lambda-1)} A_p(\nu) \tag{4.422}$$

if $n = p^{\lambda-1}\nu$ *and* $\nu \not\equiv 0 \pmod{p}$; *and*

$$A_{p^\lambda} = (p-1)p^{\lambda-1-\frac{1}{2}s\lambda} \qquad (s \equiv 0), \tag{4.4231}$$

$$A_{p^\lambda} = 0 \qquad (s \equiv 1, 3), \tag{4.4232}$$

$$A_{p^\lambda} = (-1)^{\frac{1}{2}(p-1)}(p-1)p^{\lambda-1-\frac{1}{2}s\lambda} \qquad (s \equiv 2), \tag{4.4233}$$

if $n \equiv 0 \pmod{p^\lambda}$.

If $\lambda = 2\mu + 1$, $\mu > 0$, we have $S_{h,p^\lambda} = p^\mu S_{h,p}$, by (4.132). We write

$$h = pz + h' \qquad (z = 0, 1, \cdots, p^{\lambda-1} - 1;\ h' = 1, 2, \cdots, p-1).$$

Then

$$A_{p^\lambda} = p^{-s(\mu+1)}\sum_h (S_{h,p})^s e^{-2nh\pi i/p^{2\mu+1}}$$

$$= p^{-s(\mu+1)}\sum_{h'} (S_{h',p})^s e^{-2nh'\pi i/p^{2\mu+1}} \sum_z e^{-2nz\pi i/p^{2\mu}}.$$

If $n \not\equiv 0 \pmod{p^{2\mu}}$, the sum with respect to z vanishes, and we obtain (4.421).

If $n = p^{2\mu}\nu$, we obtain

$$A_{p^\lambda} = p^{2\mu - s(\mu+1)} \sum_{h'} (S_{h',\,p})^s e^{-2\nu h'\pi i/p}.$$

If $\nu \not\equiv 0 \pmod p$, this is $p^{-(s-2)\mu} A_p(\nu)$, and we obtain (4.422). But if $\nu \equiv 0 \pmod p$ we have

$$\sum_{h'} (S_{h',\,p})^s = i^{\frac{1}{2}s(p-1)^2} p^{\frac{1}{2}s} \sum_{h'} \left(\frac{h'}{p}\right)^s,$$

and we obtain the equations (4.423).

4.43. *If λ is even and greater than 1, then*

(4.431) $$A_{p^\lambda} = 0$$

if $n \not\equiv 0 \pmod{p^{\lambda-1}}$;

(4.432) $$A_{p^\lambda} = -\,p^{\lambda-1-\frac{1}{2}s\lambda}$$

if $n = p^{\lambda-1}\nu$ *and* $\nu \not\equiv 0 \pmod p$; *and*

(4.433) $$A_{p^\lambda} = (p-1)\,p^{\lambda-1-\frac{1}{2}s\lambda}$$

if $n \equiv 0 \pmod{p^\lambda}$.

If $\lambda = 2\mu$, we have $S_{h,\,p^\lambda} = p^\mu$, by (4.133). Hence

$$A_{p^\lambda} = p^{-s\mu} \sum_{h'} e^{-2nh'\pi i/p^{2\mu}} \sum_{z} e^{-2nz\pi i/p^{2\mu-1}},$$

which is zero if $n \not\equiv 0 \pmod{p^{2\mu-1}}$, and

$$p^{2\mu-1-s\mu} \sum_{h'} e^{-2\nu h'\pi i/p}$$

if $n = p^{2\mu-1}\nu$; and the sum here is equal to -1 or $p-1$ according as ν is not or is divisible by p.

5. Summation of the singular series when $s = 8$

5. 1. The formulas of Section 4 enable us to sum the singular series whatever the value of s. I take as typical the cases $s = 8$ and $s = 5$. I suppose first that $s = 8$ and that n has no squared factor. We have to determine the factors χ_p of (4.22).

In the first place, let $p = 2$. Then, as n is not divisible by 4, we have $A_{16} = A_{32} = \cdots = 0$, by (4.321) and (4.331); and also $A_8 = 0$, by (4.321), since $\nu = 1$ and $\nu - s$ is not a multiple of 4. If n is odd, $A_4 = 0$, by (4.331); but if n is even,

$$A_4 = 2^{-3} \cos \tfrac{1}{2}\nu\pi = -\tfrac{1}{8},$$

by (4.3322). Finally $A_2 = 0$ in any case, by 4.31. Thus

(5.11) $$\chi_2 = 1 \ (n \text{ odd}), \qquad \chi_2 = \tfrac{7}{8} \ (n \text{ even}).$$

Next, suppose p odd and $p \nmid n$.[16] Then $A_{p^2} = A_{p^3} = \cdots = 0$, by (4.421) and (4.431), and $A_p = -p^{-4}$, by (4.411). Thus

$$\chi_p = 1 - p^{-4} \qquad (p \nmid n). \tag{5.12}$$

Finally suppose p odd and $p \mid n$. Then $A_{p^3} = A_{p^4} = \cdots = 0$, by (4.421) and (4.431); $A_{p^2} = p^{-7}$, by (4.432); and

$$A_p = (p-1)p^{-4},$$

by (4.415). Thus

$$\chi_p = 1 + (p-1)p^{-4} - p^{-7} = (1+p^{-3})(1-p^{-4}). \tag{5.13}$$

We have therefore

$$\begin{aligned} S &= \chi_2 \prod_{p \nmid n} (1-p^{-4}) \prod_{p \mid n} \{(1+p^{-3})(1-p^{-4})\} \\ &= \chi_2 \prod (1-p^{-4}) \prod_{p \mid n} (1+p^{-3}) = \frac{96}{\pi^4} \chi_2 \prod_{p \mid n} (1+p^{-3}), \end{aligned}$$

since

$$\prod (1-p^{-4}) = \frac{16}{15\zeta(4)} = \frac{96}{\pi^4}.$$

If n is odd,

$$\rho_8(n) = \frac{\pi^4 n^3}{\Gamma(4)} \frac{96}{\pi^4} \prod_{p \mid n} (1+p^{-3}) = 16\sigma_3(n), \tag{5.14}$$

where $\sigma_3(n)$ is the sum of the cubes of the divisors of n. If n is even,

$$\rho_8(n) = 16n^3(1-2^{-3}) \prod_{p \mid n} (1+p^{-3}) = 16\{\sigma_3'(n) - \sigma_3''(n)\}, \tag{5.15}$$

where $\sigma_3'(n)$ and $\sigma_3''(n)$ are the sums of the cubes of the even and odd divisors respectively. These are Jacobi's well-known results, proved at present, however, only when n is not divisible by any square.

5.2. Proceeding to the general case, suppose that

$$n = 2^\alpha \omega^a \omega'^{a'} \cdots \qquad (\alpha \geqq 0;\ a, a', \cdots > 0) \tag{5.21}$$

and consider first $A_{2\lambda}$.

If $\lambda = 1$, $A_{2\lambda} = 0$, by 4.31. If λ is odd and greater than 1, $A_{2\lambda} = 0$, by (4.321), unless $\nu = n/2^{\lambda-3} \equiv 0 \pmod 4$, i.e. unless $n \equiv 0 \pmod{2^{\lambda-1}}$, or unless $\lambda \leqq \alpha + 1$. If this condition is satisfied, and $n = 2^\alpha N$, so that N is odd, we have

$$A_{2\lambda} = 2^{-3(\lambda-1)} e^{-2^{\alpha+1-\lambda} N\pi i},$$

by (4.3223); and so

$$A_{2\lambda} = 2^{-3(\lambda-1)} \quad (\lambda < \alpha+1), \qquad A_{2\lambda} = -2^{-3(\lambda-1)} \quad (\lambda = \alpha+1). \tag{5.22}$$

On the other hand, if λ is even, $A_{2\lambda} = 0$, by (4.331), unless $n \equiv 0 \pmod{2^{\lambda-2}}$,

[16] Following Landau, I write $p \mid n$ for 'p is a divisor of n' and $p \nmid n$ for 'p is not a divisor of n'.

i.e. unless $\lambda \leqq \alpha + 2$. If this condition is satisfied we have, by (4.3322),

$$A_{2^\lambda} = 2^{-3(\lambda-1)} \cos(2^{a+1-\lambda} N\pi).$$

The cosine is 1 if $\lambda < \alpha + 1$, -1 if $\lambda = \alpha + 1$, and 0 if $\lambda = \alpha + 2$. Thus the equations (5.22) still hold for even values of λ. We have therefore

$$\chi_2 = 1 + 0 + 2^{-3} + 2^{-6} + \cdots + 2^{-3(a-1)} - 2^{-3a}, \tag{5.23}$$

the zero term corresponding to $\lambda = 1$.

Next suppose that p is odd. If p is not an ω, $\chi_p = 1 - p^{-4}$, as before. If $p = \omega$ and $\lambda < a + 1$, $n \equiv 0 \pmod{\omega^\lambda}$, and

$$A_{\omega^\lambda} = (\omega - 1)\omega^{-3\lambda-1},$$

by (4.415), (4.4231), or (4.433). If $\lambda = a + 1$,

$$A_{\omega^\lambda} = -\omega^{-3\lambda-1} = -\omega^{-3a-4},$$

by (4.422), (4.411), and (4.432). And if $\lambda > a + 1$,

$$A_{\omega^\lambda} = 0,$$

by (4.421) and (4.431). Thus

$$\chi_\omega = 1 + (\omega - 1)\omega^{-4} + (\omega - 1)\omega^{-7} + \cdots + (\omega - 1)\omega^{-3a-1} - \omega^{-3a-4} \tag{5.24}$$

$$= (1 - \omega^{-4})(1 + \omega^{-3} + \omega^{-6} + \cdots + \omega^{-3a}).$$

From (5.23) and (5.24) it follows, as at the end of 5.1, that

$$\rho_8(n) = 16n^3\{1 + 2^{-3} + \cdots + 2^{-3(a-1)} - 2^{-3a}\} \prod_\omega (1 + \omega^{-3} + \cdots + \omega^{-3a}),$$

it being understood that the factor in curly brackets is to be replaced by unity when $\alpha = 0$; and it is easily verified that the formulas (5.14) and (5.15) are still correct.

6. Summation of the singular series when $s = 5$ and n has no squared factors

6.1. Suppose next that $s = 5$ and that n is not divisible by any square, and first that $p = 2$. Then $A_{16} = A_{32} = \cdots = 0$, by (4.321) and (4.331). And $A_8 = 0$, by (4.321), unless $n \equiv 1 \pmod 4$, in which case, by (4.3223),

$$A_8 = 2^{-3} e^{-\frac{1}{4}(n-5)\pi i}.$$

Thus $A_8 = -\frac{1}{8}$ if $n \equiv 1 \pmod 8$, $A_8 = \frac{1}{8}$ if $n \equiv 5 \pmod 8$, and otherwise $A_8 = 0$.

Next,

$$A_4 = 2^{-3/2} \cos(\tfrac{1}{2} n\pi - \tfrac{5}{4}\pi),$$

by (4.3322), so that $A_4 = -\frac{1}{4}$ if $n \equiv 1 \pmod 4$ and $A_4 = \frac{1}{4}$ otherwise. Finally $A_2 = 0$ in all cases, by 4.31.

Collecting these results we find that

$$(6.11)\quad \chi_2 = \tfrac{5}{8}\,(n \equiv 1), \qquad \chi_2 = \tfrac{5}{4}\,(n \equiv 2, 3, 6, 7), \qquad \chi_2 = \tfrac{7}{8}\,(n \equiv 5),$$

the congruences being to modulus 8.

If p is odd and $p \nmid n$, we have

$$A_p = \left(\frac{n}{p}\right)\frac{1}{p^2}, \qquad A_{p^2} = A_{p^3} = \cdots = 0,$$

by (4.412), (4.421), and (4.431). If p is odd and $p \mid n$, we have

$$A_p = 0, \qquad A_{p^2} = -p^{-4}, \qquad A_{p^3} = A_{p^4} = \cdots = 0,$$

by (4.416), (4.432), (4.421), and (4.431). Thus

$$(6.12)\quad \chi_p = 1 + \left(\frac{n}{p}\right)\frac{1}{p^2} \quad (p \nmid n), \qquad \chi_p = 1 - \frac{1}{p^4} \quad (p \mid n).$$

If $n \not\equiv 1 \pmod 4$, we have

$$\rho_5(n) = \frac{\pi^{5/2}\, n^{3/2}}{\Gamma(5/2)}\,\frac{5}{4} \prod_{p \nmid n}\left\{1 + \left(\frac{n}{p}\right)\frac{1}{p^2}\right\} \prod_{p \mid n}\left(1 - \frac{1}{p^4}\right)$$

$$= \tfrac{5}{3}\pi^2\, n^{3/2} \prod\left(1 - \frac{1}{p^4}\right) \prod_{p \nmid n}\left\{1 - \left(\frac{n}{p}\right)\frac{1}{p^2}\right\}^{-1}.$$

Also

$$\prod\left(1 - \frac{1}{p^4}\right) = \frac{16}{15\,\zeta(4)} = \frac{96}{\pi^4}, \qquad \prod_{p \nmid n}\left\{1 - \left(\frac{n}{p}\right)\frac{1}{p^2}\right\}^{-1} = \sum\left(\frac{n}{m}\right)\frac{1}{m^2},$$

where m runs through all odd values prime to n. Hence finally

$$(6.13)\qquad \rho_5(n) = \frac{160}{\pi^2}\, n^{3/2} \sum\left(\frac{n}{m}\right)\frac{1}{m^2}.$$

If $n \equiv 1 \pmod 8$ the value of χ_2 is $\frac{5}{8}$ instead of $\frac{5}{4}$; and if $n \equiv 5 \pmod 8$ it is $\frac{7}{8}$. In these cases the numerical factor 160 must be replaced by 80 and by 112 respectively.

These are the results of Eisenstein, subsequently proved by Smith and Minkowski by means of the arithmetical theory of quadratic forms. The series (6.13) is easily summed in a finite form, by methods due to Dirichlet and to Cauchy. I have nothing to add to this part of the discussion.

7. The general case when $s = 5$

7.1. So far it has not been necessary to distinguish between one type of representation and another. At this stage the distinction between "primitive" and "imprimitive" representations becomes of importance.

A representation

$$n = x_1^2 + x_2^2 + x_3^2 + x_4^2 + x_5^2$$

is said to be *imprimitive* if x_1, x_2, x_3, x_4, x_5 possess a common factor, and *primitive* in the contrary case. It is plain that imprimitive representations can exist only when n is divisible by a square. When $s = 8$ (and the remark applies equally when s is 2, 4, or 6) the distinction is, for our purposes, irrelevant, even when n is divisible by a square: the formulas (5.14) and (5.15) are valid in any case. But when $s = 5$ the distinction is important. It will be remembered in fact, by anyone familiar with the work of Minkowski and Smith, that the right-hand side of (6.13) represents, in general, not the total number of representations but the number of primitive representations. Our series (2.26), on the other hand, gives the total number of representations; and its relation to the Smith-Minkowski series must therefore generally be more intricate than in the simplest case treated in 6.1.

The theorem which I shall prove is as follows:

The sum of the series

$$\frac{C}{\pi^2} n^{3/2} \sum \left(\frac{n}{m}\right) \frac{1}{m^2}, \tag{7.11}$$

where m runs through all odd numbers prime to n, and

$$C = 80 (n \equiv 0, 1, 4),\ C = 160 (n \equiv 2, 3, 6, 7),\ C = 112 (n \equiv 5),$$

the congruences being to modulus 8, is $\bar{r}_5(n)$, the number of primitive representations of n.

We shall require the following

LEMMA. *If*[17]

$$r(n) = \sum \phi \left(\frac{n}{q^2}\right), \tag{7.12}$$

where q^2 runs through all squared divisors of n, then

$$\phi(n) = \bar{r}(n). \tag{7.13}$$

To prove this, suppose first that n is divisible by p^2, but by no other square. Then

$$\phi(n) = r(n) - \phi\left(\frac{n}{p^2}\right) = r(n) - r\left(\frac{n}{p^2}\right) = \bar{r}(n).$$

Next, if n is divisible by p^2, p'^2, and $(pp')^2$, where $p' \neq p$, but by no other square, we have

$$\begin{aligned} \phi(n) &= r(n) - \phi\left(\frac{n}{p^2}\right) - \phi\left(\frac{n}{p'^2}\right) - \phi\left(\frac{n}{p^2 p'^2}\right) \\ &= r(n) - \bar{r}\left(\frac{n}{p^2}\right) - \bar{r}\left(\frac{n}{p'^2}\right) - r\left(\frac{n}{p^2 p'^2}\right) = \bar{r}(n). \end{aligned}$$

[17] In what follows I omit the suffix 5 in $r_5(n)$, etc.

A similar proof applies if $p = p'$; and it is plain that a repetition of the argument leads to a general proof of the lemma.

7.2. Suppose first that n is congruent to 2, 3, 6, or 7, so that n is not divisible by 4 and $\chi_2 = \frac{5}{4}$, by (6.11). If we write

$$n = 2^a N = 2^a \omega^\alpha \omega'^{\alpha'} \cdots \qquad (\alpha = 0, 1), \tag{7.21}$$

the value of χ_p requires reconsideration when p is an ω and $a > 1$. Using the formulas of 4.4, we obtain the following results.

If $a = 2b + 1$ then

$$A_\omega = 0, \quad A_{\omega^2} = (\omega - 1)\omega^{-4}, \quad A_{\omega^3} = 0, \quad A_{\omega^4} = (\omega - 1)\omega^{-7},$$
$$\cdots, \quad A_{\omega^{2b-1}} = 0, \quad A_{\omega^{2b}} = (\omega - 1)\omega^{-3b-1},$$
$$A_{\omega^{2b+1}} = 0, \quad A_{\omega^{2b+2}} = -\omega^{-3b-4}, \quad A_{\omega^{2b+3}} = A_{\omega^{2b+4}} = \cdots = 0.$$

If $a = 2b$, the values of the A's, up to $A_{\omega^{2b}}$, are as above, but

$$A_{\omega^{2b+1}} = \left(\frac{\nu}{\omega}\right)\omega^{-3b-2}, \quad A_{\omega^{2b+2}} = A_{\omega^{2b+3}} = \cdots = 0,$$

where $\nu = \omega^{-\alpha} n$. We thus find that

$$\chi_\omega = \frac{(1 - \omega^{-4})(1 - \omega^{-3b-3})}{1 - \omega^{-3}} \tag{7.22}$$

if a is odd, and

$$\chi_\omega = \frac{(1 - \omega^{-4})(1 - \omega^{-3b})}{1 - \omega^{-3}} + \left\{1 + \left(\frac{\nu}{\omega}\right)\frac{1}{\omega^2}\right\}\omega^{-3b} \tag{7.23}$$

if a is even.

Suppose now that $n = 2^a \omega^\alpha \omega'^{\alpha'} \cdots = \omega^2 d$, where d has no squared factor. Then the odd primes p fall into four classes characterized as follows.

(i) $p = \mathfrak{p}, \mathfrak{p} \nmid n$. In this case

$$\chi_p = 1 + \left(\frac{n}{\mathfrak{p}}\right)\mathfrak{p}^{-2}. \tag{7.241}$$

(ii) $p = \omega_1, \omega_1 \nmid \omega, \omega_1 | d$. In this case

$$\chi_{\omega_1} = 1 - \omega_1^{-4}. \tag{7.242}$$

(iii) $p = \omega_2, \omega_2 | \omega, \omega_2 \nmid d$. In this case a is even, say equal to $2b$, and

$$\chi_{\omega_2} = (1 - \omega_2^{-4})\left\{1 + \omega_2^{-3} + \cdots + \omega_2^{-3b+3} + \frac{\omega_2^{-3b}}{1 - \left(\frac{\nu}{\omega_2}\right)\frac{1}{\omega_2^2}}\right\}, \tag{7.243}$$

by (7.23).

(iv) $p = \omega_3, \omega_3 | \omega, \omega_3 | d$. In this case a is odd, say equal to $2b + 1$, and

$$\chi_{\omega_3} = (1 - \omega_3^{-4})(1 + \omega_3^{-3} + \cdots + \omega_3^{-3b}), \tag{7.244}$$

by (7.22). And we have

$$S = \tfrac{5}{4} \prod \chi_p = \tfrac{5}{4} \prod \chi_p \prod \chi_{\omega_1} \prod \chi_{\omega_2} \prod \chi_{\omega_3}. \tag{7.25}$$

7.3. We now multiply out the product (7.25), treating the second factors of χ_{ω_2} and χ_{ω_3} each as a sum of $b+1$ separate terms. We thus obtain

$$S = \tfrac{5}{4} \prod_p (1 - p^{-4}) \prod_p \gamma_p(n) \sum_{\lambda,\mu} \Big(\prod_{\omega_2} \theta \omega_2^{-3\lambda} \prod_{\omega_3} \omega_3^{-3\mu} \Big), \tag{7.31}$$

where

$$\gamma_p(n) = \left\{ 1 - \left(\frac{n}{p}\right) \frac{1}{p^2} \right\}^{-1}; \tag{7.311}$$

$\lambda \leqq b_2$, if $\omega_2^{a_2} = \omega_2^{2b_2}$ is the highest power of ω_2 which divides n; $\mu \leqq b_3$, if $\omega_3^{a_3} = \omega_3^{2b_3+1}$ is the highest power of ω_3 which divides n; and θ is an additional factor which is equal to 1 unless $\lambda = b_2$, and then to

$$\gamma_{\omega_2}(\nu) = \gamma_{\omega_2}(\omega_2^{-a_2} n).$$

If we denote the product which appears under the sign of summation in (7.31) by $\sigma_{\lambda,\mu}$, we have

$$\rho(n) = \frac{4\pi^2}{3} n^{3/2} S = \frac{160}{\pi^2} n^{3/2} \prod_p \gamma_p(n) \sum \sigma_{\lambda,\mu} = \sum \rho_{\lambda,\mu}, \tag{7.32}$$

say.

Suppose first that λ does not, for any ω_2, assume its maximum value b_2, so that all the θ's in $\sigma_{\lambda,\mu}$ are equal to unity; and write

$$\psi(n) = \frac{160}{\pi^2} n^{3/2} \sum \left(\frac{n}{m}\right) \frac{1}{m^2}, \tag{7.34}$$

so that

$$\psi(n) = \frac{160}{\pi^2} n^{3/2} \prod_p \gamma_p(n),$$

the product extending over all odd primes which do not divide n. Then

$$\rho_{\lambda,\mu} = \frac{160}{\pi^2} n^{3/2} \Big(\prod_{\omega_2} \omega_2^{-2\lambda} \prod_{\omega_3} \omega_3^{-2\mu} \Big)^{3/2} \prod_p \gamma_p(n) = \psi\left(\frac{n}{q^2}\right),$$

where

$$q^2 = \prod_{\omega_2} \omega_2^{-2\lambda} \prod_{\omega_3} \omega_3^{-2\mu}$$

is a typical square divisor of n, division by which does not eliminate completely any prime factor of n.

This transformation would not, as it stands, be valid if $\lambda = b_2$ for some ω_2, since there are then certain primes ω_2 which divide n and not n/q^2. But with each of these primes ω_2 there is associated an additional factor $\theta = \gamma_{\omega_2}(\nu)$ in $\sigma_{\lambda,\mu}$, and these factors provide exactly the corrective required. We have

therefore in any case $\rho_{\lambda,\mu} = \psi(n/q^2)$ and

$$r(n) = \rho(n) = \sum \psi\left(\frac{n}{q^2}\right),$$

the summation extending over all square divisors of n. And therefore, by (7.13), $\psi(n) = \bar{r}(n)$, the result required.

Our theorem is thus proved when n is congruent to 2, 3, 6, or 7 to modulus 8. In order to prove it when n is congruent to 1 or 5, we have only to write $\frac{5}{8}$ or $\frac{7}{8}$ instead of $\frac{5}{4}$ throughout our argument. It is only when n is divisible by 4 that further discussion is required.

7. 4. We have now

$$n = 2^\alpha N = 2^\alpha \omega^a \omega'^{a'} \cdots \qquad (\alpha \geqq 2).$$

The value of χ_p, when p is odd, is the same as before. The value of χ_2 may be calculated by means of the results of 4.3; and we find that

$$\chi_2 = 1 - \frac{1}{4} - \frac{1}{4.8} - \frac{1}{4.8^2} - \cdots - \frac{1}{4.8^{\beta-1}} + \frac{1}{4.8^\beta}, \tag{7.41}$$

where α is odd and equal to $2\beta + 1$ or even and equal to 2β.

Let us denote by $r^*(n)$ the number of representations of n which are *primitive so far as 2 is concerned*, that is to say in which x_1, x_2, x_3, x_4, and x_5 are not all even. It is plain that

$$r(n) = r^*(n) + r^*\left(\frac{n}{4}\right) + \cdots + r^*\left(\frac{n}{4^\beta}\right), \tag{7.42}$$

and that

$$r^*(n) = \sum \bar{r}\left(\frac{n}{q^2}\right), \tag{7.43}$$

where now the summation applies to all *odd* square divisors of n. Further, as in 7.1, we can show that *if*

$$r^*(n) = \sum \phi\left(\frac{n}{q^2}\right), \tag{7.44}$$

where the summation applies to all odd square divisors of n, then

$$\phi(n) = r(n). \tag{7.45}$$

Bearing these remarks in mind, we can complete the proof of the theorem as follows. Since $\rho(n)$, $\rho(\frac{1}{4}n)$, $\cdots$ differ only in the factor χ_2 and the outside power of n, $\frac{1}{4}n$, $\cdots$, we have, by (7.41),

$$\begin{aligned}\rho(n) - \rho\left(\frac{n}{4}\right) &= \tfrac{4}{5}\rho\left(\frac{n}{4^\beta}\right)\left\{8^\beta\left(1 - \frac{1}{4} - \cdots - \frac{1}{4.8^{\beta-1}} + \frac{1}{4.8^\beta}\right)\right. \\ &\qquad \left. - 8^{\beta-1}\left(1 - \frac{1}{4} - \cdots - \frac{1}{4.8^{\beta-2}} + \frac{1}{4.8^{\beta-1}}\right)\right\} \\ &= 4.8^{\beta-1}\rho\left(\frac{n}{4^\beta}\right).\end{aligned}$$

But, by (7.42), we have

$$r^*(n) = r(n) - r\left(\frac{n}{4}\right) = \rho(n) - \rho\left(\frac{n}{4}\right) = 4.8^{s-1}\rho\left(\frac{n}{4^s}\right) = 4.8^{s-1}r\left(\frac{n}{4^s}\right);$$

and therefore, by our previous results,

$$r^*(n) = 4.8^{s-1}\sum \bar{r}\left(\frac{n}{4^s q^2}\right) = 4.8^{s-1}\sum \psi\left(\frac{n}{4^s q^2}\right) = \tfrac{1}{2}\sum \psi\left(\frac{n}{q^2}\right),$$

the summation applying to all square divisors of $n/4^s$, or, what is the same thing, to all odd square divisors of n. Hence, by (7.45),

$$\psi(n) = 2\bar{r}(n).$$

This completes the proof of the theorem. It is easily verified that the results are in complete agreement with those of Smith.[18]

8. Concluding remarks

8.1. I have assumed, throughout this paper, that $s \leqq 8$; and it is well known that the analogous results when $s > 8$ are false.

The analysis of the paper breaks down, when $s > 8$, in one section only, namely Section 3. We can still form the singular series, and sum it by methods differing only in detail from those of Sections 4–7. We obtain a simple function of the divisors of n when s is even, a series of the Smith-Minkowski type when s is odd; and this series can still be summed in terms of the quadratic residues and non-residues of n. We can still prove, moreover, that the sum of the singular series behaves, in respect to the fundamental transformations of the modular subgroup Γ_3, exactly like the appropriate power of the theta-function ϑ, and that the function corresponding to $\eta(\tau)$ is an invariant of the group. What we cannot prove is that $\eta(\tau)$ is bounded; and the conclusion which would follow from this, namely that $\eta(\tau)$ is constant, is in fact false.

We have still, however, all the materials for a complete solution of the problem. But it is necessary to replace the analysis of Section 3 by a more complex discussion in which we deal not with a single invariant but with a linear combination of invariants, among which that represented by the sum of the singular series is the first and most important. And our conclusion will be that the number of representations of n is the sum of a function of the types considered in this paper and of a number of other arithmetical functions defined in a more recondite manner. Some of these functions have already appeared in the work of Liouville, Glaisher, and Mordell. If I do not pursue

[18] See in particular pp. 673 et seq. of the second volume of his *Collected Papers*.

this subject further, it is because such developments seem to be a part of Mr. Mordell's researches rather than of mine.

There is another question which arises more naturally out of my own researches. The singular series or principal invariant yields in any case an *asymptotic* formula for $r_s(n)$, valid without restriction on s. But, with the entry of asymptotic formulas, the peculiar interest of squares as such departs, and the problem becomes merely a somewhat trivial case of the much larger problem usually described as Waring's problem, and so of the investigations which Mr. Littlewood and I are publishing elsewhere.

A MEMOIR UPON FORMAL INVARIANCY WITH REGARD TO BINARY MODULAR TRANSFORMATIONS. INVARIANTS OF RELATIVITY*

BY

OLIVER EDMUNDS GLENN

A readable summary of existing formal concomitant theory can be given, perhaps, by referring to the scope and general trend of my former papers in this field.† The first of these gave the results of an attempt to correlate the modular and the algebraic invariant principles by emphasizing the process of transvection between a given form and the universal covariants of the modular group of binary transformations $G_{(p^2-p)(p^2-1)}$. The chief difficulty encountered was the apparent one that this process is not alone definitive for the construction of complete systems. It is also not the simplest algorism for the generation of modular concomitants. A second paper was an improvement from the latter point of view in that polar and transvectant operations characteristic of the modular theory were developed, but lack of definitiveness remained and yet remains, as each complete system which I have developed contains forms which were constructed by purely empirical methods. In subsequent papers I emphasized what seems to me at present to be the most favorable method upon which to ground a general theory by which it results that a modular covariant is a member of a finite scale of derived modular concomitants constructed by processes theoretically analogous to the important algebraical method of successive convolution.

* Presented to the Society, April 26, 1919, and December 30, 1919.

† References to other papers on this and related subjects, particularly to memoirs by Dickson, are to be found in my article in these Transactions, vol. 20 (1919), and a list of textual references for the present paper is given below. In connection with Section 8 compare my paper in the 1918 volume of the Proceedings of the National Academy of Sciences.

I. Dickson, these Transactions, vol. 12 (1911), p. 75; vol. 14 (1913), p. 299; vol. 15 (1914), p. 497. American Journal of Mathematics, vol. 31 (1909), p. 337.

II. Dickson, *The Madison Colloquium Lectures* (1913).

III. Einstein, Annalen der Physik, vol. 17 (1905).

IV. Glenn, these Transactions, vol. 20 (1919), p. 154.

V. Tolman, *The Theory of Relativity of Motion* (1917).

Fundamental likenesses between the algebraic and the modular theories are not lacking, but it seems to the writer that a theory to comprehend both effectually would not be constructed upon the principles of the algebraic symbolism but rather upon the methods such as characterize general analysis.* The present author thinks, however, that invariant theory has not yet been developed to a sufficient extent to make this synthesis feasible.

The first section of this memoir contains an introduction to a theory of modular covariants appertaining to domains of rationality. The second deals with systems of universal covariants of certain particular modular groups. The third, fourth and fifth are given over to the theory of seminvariants with application to complete systems for forms of orders 1, 2, 4, moduli 2 and 3. The remainder of the paper deals with complete systems of covariants, with regard to these particular moduli, for the general binary quantic of order four, excepting the last section which treats certain invariant problems in the theory of relativity of motion.

1. Formal covariants belonging to domains

Suppose that the coefficients of the transformation on n variables

$$\tau : x_i = \lambda_i x_1' + \mu_i x_2' + \cdots + \sigma_i x_n' \qquad (i = 1, \cdots, n),$$

are residues modulo p, a prime number, or any other quantities such that the transformation is of a finite period μ, and let the coefficients of the n-ary quantic of order m

$$f = (a_{m0\cdots0}, a_{m-1\,1\,0\cdots0}, \cdots)(x_1, \cdots, x_n)^m, \tag{1}$$

be numbers which appertain to any specific domain of rationality R. Then if τ be applied μ times in succession to f, there is produced a closed cycle of forms $f, f_1, \cdots, f_{\mu-1}$ belonging to the domain R_τ obtainable by adjunction of the fundamental quantities of the domain of the coefficients of τ to the domain R. The product

$$F_{+1} = ff_1 \cdots f_{\mu-1},$$

is an absolute covariant of τ appertaining either to R_τ or to a domain included in R_τ. The same statement is true of any symmetric function of the general type

$$\Sigma_{+1}^{(\alpha)} = \Sigma \nu f^{\alpha_0} f_1^{\alpha_1} \cdots f_{\mu-1}^{\alpha_{\mu-1}},$$

these being, of course, all rationally expressible in terms of the complete system consisting of the μ elementary symmetric functions,

$$\Sigma_{+1}, \Sigma_{+1}', \cdots, \Sigma_{+1}^{(\mu-1)}.$$

* E. H. Moore, *An Introduction to a Form of General Analysis, The New Haven Mathematical Colloquium* (1910).

If R_τ is such that one can postulate a theory of conjugacy so each form Σ_{+1} in the domain will have corresponding to it a conjugate form Σ_{-1} in the domain, then

$$P_{\pm 1} = \Sigma_{+1} \pm \Sigma_{-1} \tag{2}$$

is a covariant of τ appertaining, certainly, to a lower domain included in R_τ.

The problem of the complete system of forms P rests upon Hilbert's Lemma, since the forms P are constructed according to a law by which one can locate an arbitrarily chosen form Σ within or without the system of forms P. Thus there exists a finite set $P'_{+1}, P''_{+1}, \cdots; P'_{-1}, P''_{-1}, \cdots$ such that any other form P may be written

$$P = Q'_{+1} P'_{+1} + \cdots + Q'_{-1} P'_{-1} + Q''_{-1} P''_{-1} + \cdots,$$

and the Q's, being symmetric, are forms Σ. The conclusion follows from (2) and repetitions of the course of reasoning.

Consider the problem when the coefficients of τ are residues.

A binary system in the GF (p^2). If τ is binary $(n = 2)$ it has two poles, found from

$$\left.\begin{aligned} x_1 &\equiv \lambda_1 x_1 + \mu_1 x_2 \\ x_2 &\equiv \lambda_2 x_1 + \mu_2 x_2 \end{aligned}\right\} \pmod p,$$

and therefore roots of the quadratic congruence

$$\lambda_2 \rho^2 + (\mu_2 - \lambda_1)\rho - \mu_1 \equiv 0 \pmod p.$$

Thus the two poles are, generally, conjugate marks of the $GF(p^2)$, as

$$\rho_{\pm 1} \equiv rx \pm s \pmod p,$$

where r, s are integral and x is a galoisian imaginary. The linear forms

$$f_{\pm 1} = x_1 - \rho_{\pm 1} x_2, \tag{3}$$

are, accordingly, relative covariants of τ, universal for all transformations τ for which $\mu_2 - \lambda_1$ is congruent to a fixed residue, and appertaining to the domain R of ρ_{+1}, ρ_{-1}, in general the $GF[p^2]$.

Let us now expand the arbitrary binary form

$$f = (a_0, a_1, \cdots, a_m \mathbin{\text{\textbar\textbar}} x_1, x_2)^m$$

in terms of f_{+1}, f_{-1} as arguments, employing the substitutions inverse to (3),

$$x_1 = (-\rho_{-1} f_{+1} + \rho_{+1} f_{-1})/(\rho_{+1} - \rho_{-1}),$$

$$x_2 = (-f_{+1} + f_{-1})/(\rho_{+1} - \rho_{-1}).$$

The result is of the form

$$f \equiv \phi_m f_{+1}^m + \phi_{m-2} f_{+1}^{m-1} f_{-1} + \cdots + \phi_{-m} f_{-1}^m \pmod p. \tag{4}$$

Again if we expand f', the transformed of f by τ, in terms of the arguments

$$f'_{\pm 1} = x'_1 - \rho_{\pm 1} x'_2,$$

we get

$$f' \equiv \phi'_m f'^m_{+1} + \phi'_{m-2} f'^{m-1}_{+1} f'_{-1} + \cdots + \phi'_{-m} f'^m_{-1} \pmod p. \tag{5}$$

But $f' \equiv f$ from the equations of τ, and

$$f'_{\pm 1} \equiv \delta_{\pm 1} f_{\pm 1} \pmod p,$$

where $\delta_{\pm 1}$ are definite marks. Hence

$$f \equiv \sum_{i=0}^{m} \phi'_{m-2i} \delta^{m-i}_{+1} \delta^{i}_{-1} f^{m-i}_{+1} f^{i}_{-1} \pmod p,$$

an expansion which must be formally identical with (4) since (4) is unique. Hence the forms ϕ_{m-2i}, which are linear in $a_0, \cdots, a_m$ and which belong to the $GF(p^2)$ are invariants of τ which appertain to the domain R determined by $f_{\pm 1}$ and $\delta_{\pm 1}$. The invariant relations are

$$\phi'_{m-2i} \equiv \delta^{-m+i}_{+1} \delta^{-i}_{-1} \phi_{m-2i} \pmod p \qquad (i = 0, \cdots, m).$$

THEOREM. *The finite set ϕ_{m-2i} $(i = 0, \cdots, m)$, f_{+1}, f_{-1}, constitutes a complete system of modular concomitants of the form f, in the domain R, under the transformation τ.*

2. CERTAIN UNIVERSAL COVARIANT SYSTEMS

To determine a complete set of universal covariants of the ternary group

$$G : x \equiv x', \quad y \equiv y', \quad z \equiv x' + y' + z' \pmod p,$$

we assume a covariant in the general form

$$F = \sum_{i=0}^{m} f_{m-i} z^i,$$

where

$$f_{m-i} = a_0 x^{m-i} + a_1 x^{m-i-1} y + \cdots + a_m y^{m-i}.$$

When this is transformed into F' by G we must have $F' \equiv F \pmod p$. But

$$F' = \sum_{i=0}^{m} f_{m-i} (-x - y + z)^i \equiv F \pmod p,$$

and as the separate powers of z must accordingly have vanishing coefficients

$$\sum_{i=0}^{m} f_{m-i} (x + y)^i (-1)^i \equiv 0 \pmod p.$$

This is equivalent to the statement that F is annuled by being made simultaneous with $x + y + z = 0$. Hence

$$F \equiv (x+y+z)F_1 \pmod{p}.$$

By the theory of § 1, therefore, F is necessarily a power of the covariant

$$\psi = z[z+x+y][z+2(x+y)]\cdots[z+(p-1)(x+y)]$$
$$\equiv z^p - z(x+y)^{p-1} \pmod{p}.$$

There is a single exception to this theory, viz. when F does not contain z and hence is a binary form in the covariants x, y. Hence,

THEOREM. *A complete system of universal covariants of G is composed of x, y, ψ.*

A similar method will give the system for the group $G_1 : x \equiv x' + y'$, $y \equiv y' \pmod{p}$, previously derived,* although by a different method, by Dickson. That is, a covariant of G_1 may be assumed in the form

$$F = \sum_{i=0}^{m} a_i x^{m-i} y^i.$$

Hence

$$F' \equiv \sum_{i=0}^{m} a_i (x-y)^{m-i} y^i \equiv F \pmod{p},$$

and as the coefficients of separate powers of x must vanish identically

$$\sum_{i=0}^{m} a_i (-1)^{m-i} y^{m-i} y^i \equiv 0 \pmod{p}.$$

Hence F is annuled when made simultaneous with $x + y = 0$ and so

$$F \equiv (x+y)F_1 \pmod{p}.$$

Thus a complete system for G_1 consists of y and

$$(6) \quad \psi_1 = x(x+y)(x+2y)\cdots(x+\overline{p-1}\,y) \equiv x^p - xy^{p-1} \pmod{p}.$$

Consider the universal covariants of the simultaneous binary modular groups

$$G_1 : x_1 = x_1' + x_2', \quad x_2 = x_2'; \qquad G_2 : y_1 = y_1' + y_2', \quad y_2 = y_2'.$$

These evidently consist of

$$x_2,\ y_2,\ \psi_1 = x_1^p - x_1 x_2^{p-1}, \qquad \psi_2 = y_1^p - y_1 y_2^{p-1},$$

and other covariants which are properly simultaneous. To determine the latter we note that such a form is necessarily a homogeneous doubly binary form in the sets $x_1, x_2; y_1, y_2$, and, employing Gordan's series, we may write this quantic as follows:

*Bulletin of the American Mathematical Society, vol. 20 (1913), p. 132.

$$F = a_x^m b_y^n = a_x^m b_x^n]_{y^n} + \sum_{i=1}^{n} \frac{\binom{m}{i}\binom{n}{i}}{\binom{m+n-i+1}{i}} (xy)^i (a_x^m, b_x^n)_{y^{n-i}}^i \quad (n \leqq m).$$

Since the variables y_1, y_2 are cogredient to x_1, x_2 all terms of this expansion are covariantive. Likewise $a_x^m b_x^n$ is a covariant of G_1 and so it is expressible as a polynomial in x_2 and ψ_1. The polarizing operation indicated by $a_x^m b_x^n]_{y^n}$ gives therefore a polynomial in

$$x_2, \psi_1, \text{ and } (xy) = x_1 y_2 - x_2 y_1, y_2,$$

inasmuch as

$$\left(y_1 \frac{\partial}{\partial x_1} + y_2 \frac{\partial}{\partial x_2}\right) \psi_1 \equiv (xy) x_2^{p-2} \pmod{p}. \tag{7}$$

Every term in Σ contains (xy) as a factor whereas the other factor is a doubly binary form similar to F and of the type $\alpha_x^{m-1} \beta_y^{n-1}$ to which a similar process of reduction by expansion can be applied in turn. We have now proved the following theorem:

THEOREM. *A complete system of universal covariants of the simultaneous groups (mod p) G_1, G_2 is composed of*

$$x_2, y_2, \psi_1, \psi_2, (xy).$$

3. SEMINVARIANCY

If the form

$$f_m = (a_0, a_1, \cdots, a_m)(x_1, x_2)^m$$

be transformed by the substitution $G_t : x_1 = x_1' + t x_2', \ x_2 = x_2'$, the transformations

$$\Gamma_t : a_j' \equiv \binom{m}{j} t^j a_0 + \binom{m-1}{j-1} t^{j-1} a_1 + \cdots + a_j \quad (j = 0, \cdots, m),$$

are induced upon the coefficients.

Any homogeneous quantic in $a_0, \cdots, a_m$ which remains absolutely unaltered modulo p under Γ_t is called a seminvariant of f_m under G_t.

The principles explained in § 1 give an important class of seminvariants. For if we select for the polynomial f any one of the N possible linear quantics in the a's

$$f = \alpha_0 a_0 + \alpha_1 a_1 + \cdots + \alpha_m a_m,$$

where α_i $(i = 0, \cdots, m)$ takes any value $0, 1, \cdots, p-1$, and where the α's are selected so the different f forms are essentially different modulo p, then, if f is not itself a seminvariant, the successive application of the transformation Γ_1 of period p, to f, gives p forms

$$f, f_1, \cdots, f_{p-1},$$

and any elementary symmetric function

$$S_r = \Sigma \nu f f_1 \cdots f_r \qquad (r = 0 \text{ to } p-1), \tag{8}$$

which is not congruent to zero is a seminvariant. Some of these functions S will be shown to be reducible in terms of other seminvariants but they must all be taken into account in the problem of the construction of a complete seminvariant system.

4. An algorism for the construction of seminvariants, invariants, and covariants

If

$$K_\nu = C_0 x_1^\nu + C_1 x_1^{\nu-1} x_2 + \cdots + C_\nu x_2^\nu$$

is any formal covariant modulo p of f_m the copied form constructed from the coefficients C_i upon the model of S_r of the preceding section is a seminvariant of f_m.

Various methods of constructing covariants are available. Polarization of K_ν by the modular operators

$$w = x_1^p \frac{\partial}{\partial x_1} + x_2^p \frac{\partial}{\partial x_2},$$

$$E = a_0^p \frac{\partial}{\partial a_0} + a_1^p \frac{\partial}{\partial a_1} + \cdots + a_m^p \frac{\partial}{\partial a_m},$$

produces covariants if $\nu \not\equiv 0 \pmod p$. Substitution in the conjugate to a seminvariant $S(a_0, \cdots)$ under the permutations $s = (a_0 a_m)(a_1 a_{m-1}) \cdots$, according to Γ_t and reduction of the powers of the residue t by Fermat's theorem gives a covariant in non-homogeneous form. It can be made homogeneous by the replacement $t = x_1/x_2$. This process yields a covariant only on condition that S satisfies a definite pair of necessary and sufficient conditions as given in one of my former papers. When this principle is applied in the case of the seminvariant leading coefficient of K_ν where ν is a number of the form

$$\nu = (sp - \nu')(p-1) = \sigma(p-1)$$

there results a covariant

$$[K_\nu] = [C_0 + (K_\nu)] x_1^{p-1} + \sum_{r=0}^{p-2} \sum_{s=0}^{\sigma-1} C_{\nu - s(p-1) - r} x_1^r x_2^{p-r-1}, \tag{9}$$

where (K_ν) is the pure invariant

$$(K_\nu) = C_{p-1} + C_{2(p-1)} + \cdots + C_{(\sigma-1)(p-1)}. \tag{10}$$

If we form the product of any two covariants

$$F_{m'} = (C_0, C_1, \cdots, C_{m'})(x_1, x_2)^{m'}, \qquad G_n = (D_0, D_1, \cdots, D_n)(x_1, x_2)^n$$

in which $m' + n$ is a number of the form $\sigma(p-1)$ and construct the formulas analogous to (9), (10) for the result we obtain the concomitants of f_m,

$$(F_{m'} G_n) = \sum_{i=0}^{j(p-1)} \sum_{j=1}^{\sigma-1} C_i D_{j(p-1)-i}, \tag{11}$$

$$[F_{m'} G_n] = [C_0 D_0 + (F_{m'} G_n)] x_1^{p-1} + \sum_{r=0}^{p-2} \sum_{s=0}^{\sigma-1} \sum_{i=0}^{t} C_i D_{(\sigma-s)(p-1)-r-i} x_1^r x_2^{p-1-r} \quad \{t = (\sigma - s)(p-1) - r\}. \tag{12}$$

In connection with a single covariant K_ν we designate the derived set of $\mu = p + \nu + 2$ concomitants

$$(K_\nu),\ w^\lambda[K_\nu] \quad (\lambda = 0, 1, \cdots, p-1), \qquad w^t K_\nu \quad (t = 0, 1, \cdots, \nu), \tag{13}$$

as the *μ-adic scale* for K_ν.

The processes (11), (12), (13) constitute the methods to which we referred in the introduction as being theoretically analogous to symbolical convolution in the algebraic invariant theory.*

5. Seminvariant systems

Complete system modulo 2 of the set f_1, f_2. The algorism described above relates in its complete generality to simultaneous systems. Consider the typical problem of the seminvariant system (mod 2) of a linear form and a quadratic,

$$f_1 = a_0 x_1 + a_1 x_2, \qquad f_2 = b_0 x_1^2 + b_1 x_1 x_2 + b_2 x_2^2.$$

The induced transformations Γ_t under G_t are these;

$$a_0' = a_0, \qquad a_1' = a_0 t + a_1,$$

$$b_0' = b_0, \qquad b_1' = b_1, \qquad b_2' = b_0 t^2 + b_1 t + b_2,$$

and, if we put $b_0 = 0$, G_1, Γ_1 become

$$a_1 = a_1' + a_0'; \qquad b_2 = b_2' + b_1',$$

$$a_0 = a_0', \qquad b_1 = b_1',$$

and these are the transformations of the simultaneous groups G_1, G_2 of § 2. Hence, if

$$S(b_0, b_1, b_2, a_0, a_1)$$

is any seminvariant of the system f_1, f_2,

$$S(0, b_1, b_2, a_0, a_1) \equiv F(a_0, b_1, \psi_1, \psi_2, C) \pmod 2,$$

* The notation here adopted was first given in the Proceedings of the National Academy of Sciences, vol. 5 (1919), pp. 107–110.

where F is a polynomial in its arguments and

$$\psi_1 = a_1^2 + a_0 a_1, \qquad \psi_2 = b_2^2 + b_1 b_2, \qquad C = a_0 b_2 + a_1 b_1.$$

However the expressions ψ_2, C are merely fragments of seminvariants which remain after one makes $b_0 = 0$. We now require these in their entirety. Apply the transformation of period 2

$$b_0 = b_0', \qquad b_1 = b_1', \qquad b_2 = b_0' + b_1' + b_2'$$

to $f = b_2$ and we get $f' = b_0 + b_1 + b_2$. The sum $f + f'$ is the reducible seminvariant $b_0 + b_1$ whereas

$$s = ff' = (b_0 + b_1 + b_2) b_2 = \psi_2 + b_0 b_2.$$

Again the seminvariant leading coefficient of $[f_1 f_2]$ is $a_0 b_0 + (f_1 f_2)$, where $(f_1 f_2)$ is the invariant

$$(14) \qquad (f_1 f_2) = a_0 (b_1 + b_2) + a_1 (b_0 + b_1) = a_0 b_1 + C + b_0 a_1$$
$$= a_0 b_1 + \rho.$$

Therefore

$$S(b_0, b_1, b_2, a_0, a_1) = F(a_0, b_0, b_1, \psi_1, s, \rho) + b_0 F_1(a_0, a_1, b_0, b_1, b_2).$$

But F_1 is a seminvariant and can be reduced in the form

$$F_1 = \Phi(a_0, b_0, b_1, \psi_1, s, \rho) + b_0 \Phi_1(a_0, a_1, b_0, b_1, b_2),$$

and the process can be repeated. Hence,

THEOREM. *A complete system of seminvariants of the set f_1, f_2, modulo 2, consists of*

$$a_0, \ b_0, \ b_1, \ \psi_1, \ s, \ \rho.$$

The expressions $k = b_0 s$, $q = s + b_0 b_1 + b_0^2$ are invariants, as are

$$Q' = a_0^2 + \psi_1, \qquad L' = a_0 \psi_1,$$
$$E(f_1 f_2) = a_0 (b_1^2 + b_2^2) + a_1 (b_0^2 + b_1^2).$$

We may note the following syzygies:

$$(15) \qquad \begin{aligned} &\rho^2 + (a_0 b_0 + a_0 b_1)\rho + a_0^2 s + (b_0^2 + b_1^2)\psi_1 \equiv 0, \\ &a_0^2 + \psi_1 + Q' \equiv 0, \\ &b_0^3 + b_0^2 b_1 + b_0 q + k \equiv 0, \\ &a_0^3 + L' + a_0 Q' \equiv 0, \end{aligned}$$

and we now see that the general seminvariant S can be reduced, by means of these congruences, to the finite form

$$S \equiv \phi_0 + \phi_1 b_0 + \phi_2 b_0^2 + a_0(\psi_0 + \psi_1 b_0 + \psi_2 b_0^2) + a_0^2(\chi_0 + \chi_1 b_0 + \chi_2 b_0^2), \tag{16}$$

where ϕ_i, ψ_i, χ_i are polynomials in the pure invariants $q, k, b_1, L', Q', (f_1 f_2)$.

A complete system modulo 3 *of* f_4. I shall refer informally at this juncture to an important fact which is to be employed later on in the paper viz. that every form f_m of order $> p^2 - 1$ is reducible modulo p in terms of the universal covariants* of $G_{(p^2-p)(p^2-1)}$ i.e.

$$L = x_1^p x_2 - x_1 x_2^p, \qquad Q = (x_1^{p^2} x_2 - x_1 x_2^{p^2})/L, \tag{17}$$

and covariants of f_m. Thus we need to consider in detail only those covariants of order $< p^2$.

For the sake of constructing seminvariants by the principle of copied forms we first derive those of the first degree (modulo 3) for the quantics of orders $2, 4, 6, 8 = 3^2 - 1$.

The linear seminvariant of f_2 mod 3 is a_0; those of a quartic $f_4 = a_0 x_1^4 + \cdots$ may be found by applying the induced transformations Γ_1 to the general expression

$$S = \alpha a_0 + \beta a_1 + \gamma a_2 + \delta a_3 + \epsilon a_4,$$

assumed to be a seminvariant, making $S' \equiv S \pmod 3$ and solving for $\alpha, \beta, \gamma, \delta, \epsilon$. The induced transformations are

$$\begin{gathered} a_0' = a_0, \qquad a_1' = a_0 + a_1, \qquad a_2' = a_2, \\ a_3' = a_0 + 2a_2 + a_3, \qquad a_4' = a_0 + a_1 + a_2 + a_3 + a_4, \end{gathered} \tag{18}$$

and S reduces to $\alpha a_0 + \gamma a_2$. Hence the linear seminvariants are

$$a_0, \ a_2 = (f_4). \tag{19}$$

For f_6 the induced transformations are

$$a_0' = a_0, \quad a_1' = a_1, \quad a_2' = 2a_1 + a_2, \quad a_3' = 2a_0 + a_1 + a_2 + a_3,$$
$$a_4' = a_1 + a_4, \quad a_5' = 2a_1 + a_2 + 2a_4 + a_5,$$
$$a_6' = a_0 + a_1 + a_2 + a_3 + a_4 + a_5 + a_6,$$

and the linear seminvariants prove to be

$$a_0, \ a_1, \ a_2 + a_4 = (f_6). \tag{20}$$

In the case of the octic f_8 we find, for Γ_1,

$$a_0' = a_0, \quad a_1' = 2a_0 + a_1, \quad a_2' = a_0 + a_1 + a_2, \quad a_3' = 2a_0 + a_3,$$
$$a_4' = a_0 + 2a_1 + 2a_3 + a_4, \quad a_5' = 2a_0 + 2a_1 + 2a_2 + a_3 + a_4 + a_5,$$
$$a_6' = a_0 + a_3 + a_6, \quad a_7' = 2a_0 + a_1 + 2a_3 + a_4 + 2a_6 + a_7,$$
$$a_8' = a_0 + a_1 + a_2 + a_3 + a_4 + a_5 + a_6 + a_7 + a_8,$$

* Compare II.

and for the seminvariants

$$a_0, \qquad V = a_1 - a_3, \qquad a_2 + a_4 + a_6 = (f_8).$$

Seminvariants of the second and third degrees of f_4. The algebraic invariants i, J, and the hessian, H, of f_4, which would also be modular concomitants, are reducible modulo 3; $H \equiv 2a_2 f_4 \pmod 3$.

The seminvariant constructed on the model of V for the octic f_4^2 is

$$P = a_0 a_1 + 2a_1 a_2 + 2a_0 a_3,$$

while

$$q = (f_4^2) = a_1^2 + a_2^2 + a_3^2 + 2a_0 a_2 + 2a_0 a_4 + 2a_1 a_3 + 2a_2 a_4.$$

The transformations of (18) which affect a_0, $-a_2$, a_3 but not a_1 or a_4, form the particular ternary group G of § 2 and hence the following quantic is a seminvariant

$$N = a_3^3 - (a_0^2 + a_0 a_2 + a_2^2) a_3.$$

If we take $f = a_1$ in the theory of § 1, then $f' = a_0 + a_1$, $f'' = a_0 - a_1$ (by (18)), from which we obtain

$$\beta = a_1^3 - a_0^2 a_1 = -ff'f'',$$

or this seminvariant is obtained as a case of the universal covariant ψ_1 of the group G_1 (cf. (6)).

Selection of a_4 as f and application of (18) gives

$$\gamma_0 = a_4(a_0 + a_1 + a_2 + a_3 + a_4)(a_0 - a_1 + a_2 - a_3 + a_4) = ff'f'',$$

whereas, if we take $f = a_0 + a_4$ this method gives the invariant

$$X = (a_0 + a_4)(a_0 + a_1 - a_2 + a_3 - a_4)(a_0 - a_1 - a_2 - a_3 - a_4) + a_0^2 a_2.$$

Lastly, if we select $f = a_1 + a_4$ so that $-f'' = a_2 - a_3 + a_4$,

$$f' = a_0 + a_1 - a_2 - a_3 - a_4,$$

the seminvariant $-ff'f''$ is congruent to $2\gamma_0 + D + H + a_2 P$, where

$$H = a_0^2 a_4 + a_1^2 a_2 - a_0 a_1 a_3,$$

$$D = a_0 a_2 a_3 - a_1 a_2 a_4 - a_1^2 a_3 + a_1 a_3^2 + a_0 a_1 a_4 - a_0 a_3 a_4.$$

We shall show that γ_0 can be eliminated by means of a syzygy between a_0, γ_0, q, X. Note that D, β, P are skew. I desire to enunciate and to treat at some length the following

THEOREM. *The nine seminvariants* β, D, H, N, P, q, X, a_0, $a_2 = (f_4)$ *compose a complete system for* f_4 *modulo* 3.

Suppose that

$$S = S(a_0, a_1, a_2, a_3, a_4)$$

is an arbitrary seminvariant modulo 3 of f_4. Then it will ordinarily be possible to separate S into two seminvariantive portions

$$S = S_1 + S_2$$

where S_2 contains a_4 but S_1 does not. Since S_1 is not affected by the fifth congruence of (18) we write

$$S_1 = S_1(a_0, a_1, a_2, a_3)$$

and note that S_1' is invariantive under the binary groups G_1, G_2 of § 2, i.e.

$$\begin{aligned} &G_1 : a_0' = a_0, \qquad a_1' = a_0 + a_1, \\ &G_2 : a_0' - a_2' = a_0 - a_2, \qquad a_3' = a_0 - a_2 + a_3, \end{aligned} \tag{21}$$

therefore S_1 is a polynomial in a_0, a_2, β, N, P, since

$$-P = \begin{vmatrix} a_0 & a_1 \\ a_0 - a_2 & a_3 \end{vmatrix}.$$

With reference to S_2 we next assume that $a_0 = a_2 = 0$, momentarily, and note that the group (18) is then equivalent to G (§ 2). Thus $S_2(0, a_1, 0, a_3, a_4)$ is a polynomial in a_1, a_3 and

$$\gamma_0' = a_1^3 - (a_1 + a_3)^2 a_4 = \gamma_0]_{a_0=a_2=0}.$$

The coefficients of the different powers of γ_0' in this expression are polynomials in a_1, a_3 and of such a nature as to be expressible in terms of functions of a_1, a_3 which are fragments of seminvariants, which remain when we make $a_0 = a_2 = 0$. Four seminvariants which reduce to expressions in a_1, a_3 when $a_0 = a_2 = 0$ are β, q, D, N and we here make an assumption,* to be treated in a later paragraph from another point of view, viz., that the coefficients of the different powers of γ_0' in the expression for $S_2(0, a_1, 0, a_3, a_4)$ are all rationally expressible in terms of

$$\beta_1 = a_1^3, \qquad N_1 = a_3^3, \qquad q_1 = a_1^2 + 2a_1 a_3 + a_3^2, \qquad D_1 = a_1 a_3^2 - a_1^2 a_3, \tag{22}$$

these being the only ones of the nine seminvariants which reduce to expressions in a_1, a_3 when $a_0 = a_2 = 0$ and $a_4 \neq 0$. Under this assumption we now have

$$S_2(a_0, a_1, a_2, a_3, a_4) \equiv \phi(\beta, D, N, q, \gamma_0) + a_0 \phi_0 + a_2 \phi_2 \pmod 3, \tag{23}$$

where ϕ_0, ϕ_2 are polynomials in a_0, a_1, a_2, a_3, a_4, not both free from a_4, and such that $a_0 \phi_0 + a_2 \phi_2$ is a seminvariant. If, then, ϕ_0, ϕ_4 are both seminvariants they are of the type of the original S and the reductions can be applied to them in turn. If they are not separately seminvariantive

$$\psi = a_0 \phi_0 + a_2 \phi_2$$

is a seminvariant every term of which contains either a_0 or a_2 as a factor and which is not free from a_4. Moreover, under (18),

$$\psi' = a_0(\phi_0 + \delta\phi_0) + a_2(\phi_2 + \delta\phi_2) \equiv \psi,$$

where the expressions $\delta\phi_i$ $(i = 0, 2)$ are increments. Hence

$$\frac{\delta\phi_0}{a_2} = -\frac{\delta\phi_2}{a_0}, \tag{24}$$

i.e. $\delta\phi_0$ contains the factor a_2 and hence ϕ_0 is a function of both a_3, a_4 combined in the form of a difference, or of $a_1 - a_3$, whereas ϕ_2, in order that $\delta\phi_2$ may contain a_0 as a factor, must be a function of a_1 only or contain a power of a_1 as a factor. It follows, since

$$H = \begin{vmatrix} a_0 & -a_2 \\ a_1^2 & a_0 a_4 - a_1 a_3 \end{vmatrix},$$

that*

$$\psi = ZM,$$

where Z is a polynomial in H, P with numerical coefficients, whereas M is a seminvariant of arbitrary form.

The reader has observed, however, that this proof is definitely intuitional at two points, i.e. at the conclusions which are starred. To reduce a possible doubt as to these inferences to a minimum I set for myself the task of computing and reducing in terms of the nine fundamental seminvariants the 285 seminvariants of f_4 modulo 3 constructed as elementary symmetric functions of f, f', f'' where f is a chosen linear expression of the form

$$f = \alpha_0 a_0 + \alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3 + \alpha_4 a_4,$$

the α's being the numbers 0, 1, 2 in definite sequence, and f', f'' the two functions transformed of f by (18) (compare § 3). This problem has a certain ulterior significance as it is not impossible that a theory of complete systems might be formulated from the point of view of the symmetric function theory.§ I have no suggestions to make in this connection at present, however; the point here being that all of the 285 seminvariants so calculated proved to be reducible in terms of β, D, H, N, P, q, X, a_0, a_2.

It would not be relevant to reproduce the computations for these various cases, but I give the results for two types f which are typical.

$$f = a_3 + 2a_4, \quad f' = 2a_1 + a_2 + 2a_4, \quad f'' = a_0 + a_1 + 2a_3 + 2a_4. \tag{25}$$

$$\Sigma f = a_0 + a_2, \quad \Sigma ff' = 2(q + P) + a_2^2,$$

$$\Sigma ff'f'' = 2\gamma_0 + 2a_2 q + Ha_2^3 - a_0 a_2^2 + D.$$

§ Compare the third reference of I.

$$(26)\ f = a_0 + a_3 + a_4, \qquad f' = a_1 + 2a_3 + a_4, \qquad f'' = a_0 + 2a_1 + 2a_2 + a_4.$$

$$\Sigma f = 2(a_0 + a_2), \qquad \Sigma ff' = a_0^2 + 2q + P + a_0 a_2 + a_2^2,$$

$$\Sigma ff' f'' = D + a_2 P + 2H + \gamma_0 + (a_0 + 2a_2)(P + 2q) + 2a_0 a_2 (a_0 + a_2).$$

Congruential syzygies. If the course of development similar to that followed in § 5 is to be followed it is now required to derive a sufficient number of syzygies which can be employed as reduction moduli to reduce the general seminvariant S to a polynomial *in finite form* (cf. (16)) in the nine seminvariants of the complete system. This again is a formal problem of much difficulty, but by simple combinations we get the two below, which one can easily verify:

$$(27)\quad \left.\begin{aligned} a_0^3 + 2a_0 q + \gamma_0 + 2X + 2a_0 a_2 (a_0 + 2a_2) &\equiv 0 \\ a_0^4 + 2a_0^2 q + 2a_0 X + k + 2a_0^2 a_2 (a_0 + 2a_2) &\equiv 0 \end{aligned}\right\} \pmod 3,$$

where k is the invariant $a_0 \gamma_0$.

We next represent by S' the result of placing $a_1 = 0$ in a seminvariant S and by S'' the result of substituting zero for both a_1 and a_3 in S. Then

$$q'' = (a_0 + a_2 + a_4)^2 + 2a_0^2 + 2a_4^2,$$

$$\gamma_0'' = (a_0 + a_2 + a_4)^2 a_4,$$

$$R'' = (a_0 - a_2 - a_4)(a_2 + a_4) a_4,$$

where

$$R = (a_1 + a_4)(a_2 - a_3 + a_4)(a_0 + a_1 - a_2 - a_3 - a_4) = -ff' f'',$$

then

$$R'' + \gamma_0'' = a_0^3 a_4,$$

$$2a_0^2 (q'' + a_2^2) = a_0^3 a_2^2 + a_0^3 a_2 + a_0^3 a_4 + a_0^2 a_2 a_4,$$

hence

$$2a_0^2 (q'' + a_2^2) = a_0^2 a_2^2 + a_0^3 a_2 + (a_0 + a_2)(R'' + \gamma_0''),$$

and, in consequence,

$$\Theta' = a_0^3 a_2 + (a_0 + a_2)(R' + \gamma_0') + a_0^2 (q' + 2a_2^2)$$

is a function which vanishes when we place $a_3 = 0$.

Furthermore

$$q' = a_2^2 + a_3^2 + 2a_0 a_2 + 2a_0 a_4 + 2a_2 a_4,$$

$$P' = 2a_0 a_3,$$

$$N' = a_3^3 - a_2^2 a_3 - (a_0^2 + a_0 a_2) a_3,$$

$$= a_3^3 - a_2^2 a_3 + (a_0 + a_2) P',$$

$$\gamma_0' = a_0^2 a_4 + a_2^2 a_4 + 2a_0 a_2 a_4 + 2a_0 a_4^2 + 2a_2 a_4^2 + 2a_3^2 a_4 + a_4^3,$$

$$R' = a_0 a_2 a_4 - a_0 a_3 a_4 + a_0 a_4^2 - a_2^2 a_4 + a_2 a_4^2 + a_3^2 a_4 - a_4^3,$$

and thus we get

$$\Theta' \equiv 2a_0 a_3 [a_0 a_4 + 2a_0 a_3 + a_2 a_4]$$
$$\equiv 2P'[q' + 2P' + a_0 a_2 + a_2^2] + 2a_0 N' + (a_0^2 + a_0 a_2) P'.$$

Hence a function which vanishes when $a_1 = 0$ is

$$\Theta = P^2 + 2(q + a_0 a_2 + a_2^2) P + 2a_0 N + (a_0^2 + a_0 a_2) P$$
$$+ 2a_0^2 (q + 2a_2^2) + 2a_0^3 a_2 + 2(a_0 + a_2)(R + \gamma_0),$$

and so Θ is congruent to a seminvariant which contains β as a factor. We find by actual evaluation $\Theta = (2a_0 + a_2)\beta$. Also the seminvariant R is reducible in the form

$$R \equiv 2\gamma_0 + H + D + a_2 P \pmod 3,$$

therefore we have derived the following syzygy of degree 4:

$$P^2 + 2(q + 2a_0^2 + a_2^2) P + 2a_0 N + 2a_0^2 (q + 2a_2^2)$$
$$+ 2a_0^3 a_2 + 2(a_0 + a_2)(D + a_2 P + H) + (a_0 + 2a_2)\beta = 0. \tag{28}$$

Some of the seminvariants involved are of even and others of odd weight, hence the relation is equivalent to two syzygies obtained by equating the terms of even and odd weights to zero separately. Thus, finally, we get

$$P^2 + 2a_0^2 (q + 2a_2^2) + 2a_0^3 a_2 + 2(a_0 + a_2) H = 0,$$
$$P(q + 2a_0^2 + a_2^2) + a_0 N + (a_0 + a_2)(D + a_2 P) + (2a_0 + a_2)\beta = 0. \tag{29}$$

The seminvariant N is connected with a third degree invariant

$$\tau = a_3^3 - a_1^3 + a_1 a_2^2 - a_2^2 a_3,$$

by the linear relation

$$N = \tau + \beta + (a_0 + a_2) P,$$

and we therefore see that the equivalents of five of the nine seminvariants, i.e., a_2, q, X, D, N, are irreducible invariants. The arbitrary seminvariant S can be reduced, as a result of this fact, by the relations (27) and the first syzygy of (29), to a polynomial of indefinite order in two seminvariants β, H with coefficients which are of degree three in a_0 and linear in P, in the form

$$S = \sum_{i=1}^{\nu} [\Theta_{0i} + \Theta_{1i} a_0 + \Theta_{2i} a_0^2 + \Theta_{3i} a_0^3$$
$$+ (\Delta_{0i} + \Delta_{1i} a_0 + \Delta_{2i} a_0^2 + \Delta_{3i} a_0^3) P] \beta^{\tau_i} H^{\sigma_i}, \tag{30}$$

where $\tau_1 = \sigma_1 = 0$, and θ_{hi} and Δ_{hi} $(i = 1, \cdots, \nu;\ h = 0, \cdots, 3)$ are polynomials with numerical coefficients in the pure invariants

$$a_2, D, q, \tau, X, k. \tag{30$_1$}$$

There doubtless exist two syzygies, one headed by a low power of each seminvariant β, H, of degree higher than four, and involving additional members of a complete system of pure invariants, by means of which S could be reduced to a form finite and numerical as to powers of β, H also.

6. Covariant systems of the set f_1, f_2, and for the binary quartic f_4, modulo 2

I shall prove that a fundamental system of covariants of f_4 modulo 2 is furnished by construction of a simultaneous system of a certain pair of covariants of f_4, one a quadratic and the other linear. For this reason I select for the next topic to be treated the problem of the covariant system modulo 2 of the set f_1, f_2 (cf. § 5).

The formula (16) for the arbitrary seminvariant leading coefficient of a simultaneous covariant of f_1, f_2 suggests the form of the covariants requisite to reduce the simultaneous covariant of general type. That is we require covariants of the various possible orders led by $a_0^i b_0^j$ $(i, j = 0, 1, 2)$ and by invariants.

When the modulus is 2 the concomitant scale (13) of § 4 is not complete without the addition of the two forms*

$$\{K_\nu\} = C_0 x_1^2 + (K_\nu) x_1 x_2 + C_\nu x_2^2,$$

$$\{\overline{K}_\nu\} = C_0 x_1^3 + I_1 x_1^2 x_2 + I_2 x_1 x_2^2 + C_\nu x_2^3,$$

where I_1, I_2 are definite linear expressions in $C_1, \cdots, C_{\nu-1}$ such that

$$I_1 + I_2 \equiv (K_\nu) \pmod 2.$$

There exist no covariants of f_1 led by invariants, no quadratic covariant of any system of forms (mod 2) led by an invariant I, other than IQ, and the only possible invariant leading coefficient for a covariant C of the set f_1, f_2, which is a polynomial in q, k, b_1, L', Q', contains $b_1 q + k$ as a factor. The first two of these statements are easily verified; it will be sufficient to prove the latter for a cubic covariant in view of the form of $\{\overline{K}_\nu\}$. If I is the assumed leading coefficient it can be reduced, by virtue of the congruence

$$b_0 q = b_0(s + b_0 b_1 + b_0^2) = b_0^3 + b_0^2 b_1 + k, \tag{31}$$

to the form

$$I = \psi_0 + \psi_1 b_0 + \psi_2 b_0^2$$

where ψ_0, ψ_1, ψ_2 are seminvariants. Apply the induced transformations under $x_2 = x_1' + x_2'$, $x_1 = x_1'$ to

$$C = I x_1^3 + S_1 x_1^2 x_2 + S_2 x_1 x_2^2 + (\psi_0 + \psi_1 b_0 + \psi_2 b_0^2) x_2^3$$

* Compare IV.

and equate the result to C and there results

$$\delta S_1 = \delta S_2 = I, \qquad \delta S_1 + \delta S_2 + S_1 + S_2 + I = 0,$$

$$I = \delta S_1 + \delta S_2 + S_1 + S_2 + I + \psi_0 + (b_0 + b_1 + b_2)\psi_1 + (b_0 + b_1 + b_2)^2 \psi_2. \tag{32}$$

Hence

$$I = S_1 + S_2 = (b_1 + b_2)[\psi_1 + (b_1 + b_2)\psi_2].$$

But an invariant with the factor $b_1 + b_2$ has the factor

$$b_1 q + k = (b_0 + b_1)(b_0 + b_2)(b_1 + b_2), \tag{33}$$

which was to be proved.

The following cubic covariant which I constructed by empirical methods is led by $b_1 q + k$ and the product of it by an arbitrary invariant of f_1, f_2 is the general cubic with an invariant leader of the stated type:

$$\begin{aligned} C = {} & (b_0 + b_1)(b_0 + b_2)(b_1 + b_2)x_1^3 \\ & + (b_0^3 + b_1^3 + b_2^3 + b_0 b_1^2 + b_1 b_2^2 + b_0^2 b_2)x_1^2 x_2 \\ & + (b_0^3 + b_1^3 + b_2^3 + b_0^2 b_1 + b_1^2 b_2 + b_0 b_2^2)x_1 x_2^2 \\ & + (b_0 + b_1)(b_0 + b_2)(b_1 + b_2)x_2^3. \end{aligned}$$

The form C is reducible in terms of concomitants from the sets (34), (35), (36).

An invariant leading coefficient which is properly simultaneous and therefore not merely a polynomial in q, k, b_1, L', Q' must contain $(f_1 f_2)$ as a factor, this being the only remaining invariant in the set from which ϕ_0 is formed. We shall construct a cubic covariant B led by $(f_1 f_2)$. The most general form of ϕ_0 is

$$\phi_0 = \lambda + (f_1 f_2)\mu,$$

where λ is a polynomial in q, k, b_1, L', Q' and μ is an invariant. Assuming that ϕ_0 leads a cubic covariant D', $D' - \mu B$ is led by λ and so $\lambda = (b_1 q + k)\nu$ where ν is an invariant. Thus

$$D' = \nu C + \mu B.$$

I now give a list of cubic covariants needed, with the seminvariant (or invariant) leader juxtaposed.

$$\begin{gathered} a_0,\ Qf_1; \qquad a_0^2,\ QEf_1; \qquad b_0 + b_1,\ Q[f_2]; \\ b_0^2 + b_1^2,\ Q[Ef_2]; \qquad a_0 b_0,\ f_1 f_2; \qquad a_0 b_0^2,\ f_1 Ef_2; \\ a_0^2 b_0,\ f_2 Ef_1; \qquad a_0^2 b_0^2,\ Ef_1 Ef_2; \qquad (f_1 f_2),\ Q[f_1 f_2] + f_1 f_2 = B. \end{gathered} \tag{34}$$

The quadratic and linear covariants required are given in the following lists respectively:

$$(35)\qquad \begin{aligned} &a_0,\ wf_1;\quad a_0^2,\ f_1^2;\quad b_0,\ f_2;\quad b_0^2,\ Ef_2;\quad a_0 b_0,\ \{f_1 f_2\};\\ &a_0^2 b_0,\ \{Ef_1 \cdot f_2\};\quad a_0 b_0^2,\ \{f_1\, Ef_2\};\quad a_0^2 b_0^2,\ \{Ef_1\, Ef_2\}. \end{aligned}$$

$$(36)\qquad \begin{aligned} &a_0,\ f_1;\quad a_0^2,\ Ef_1;\quad b_0 + b_1,\ [f_2];\quad b_0^2 + b_1^2,\ E[f_2];\\ &a_0 b_0 + (f_1 f_2),\ [f_1 f_2];\quad a_0 b_0^2 + (f_1\, Ef_2),\ [f_1\, Ef_2];\\ &a_0^2 b_0 + (Ef_1 \cdot f_2),\ [Ef_1 f_2];\quad a_0^2 b_0^2 + (Ef_1\, Ef_2),\ [Ef_1\, Ef_2]. \end{aligned}$$

Observe next that the cubic covariant of the set f_1, f_2 of arbitrary form, i.e. (see (16))

$$\begin{aligned} K = Sx_1^3 + \cdots = [\,&\phi_0 + b_1\phi_1 + b_1^2\phi_2 + \phi_1(b_0 + b_1)\\ &+ \phi_2(b_0^2 + b_1^2) + a_0(\psi_0 + \cdots) + \cdots]\,x_1^3 + \cdots \end{aligned}$$

is reducible in the form

$$(37)\qquad \begin{aligned} K \equiv\ &\mu B + \nu C + \phi_1 Q[f_2] + \phi_2 Q[Ef_2] + \psi_0 Qf_1 + \psi_1 f_1 f_2\\ &+ \psi_2 f_1 Ef_2 + \chi_0 QEf_1 + \chi_1 f_2\, Ef_1 + \chi_2\, Ef_1\, Ef_2 + K'L \pmod 2, \end{aligned}$$

where K' is an invariant.

Moreover the arbitrary quadratic is reduced by the formula ($\phi_0 = 0$)

$$(38)\qquad \begin{aligned} K = \phi_1 f_2 + \phi_2\, Ef_2 + \psi_0\, wf_1 + \psi_1\{f_1 f_2\} + \psi_2\{f_1\, Ef_2\}\\ + \chi_0 f_1^2 + \chi_1\{Ef_1 f_2\} + \chi_2\{Ef_1\, Ef_2\}, \pmod 2. \end{aligned}$$

As a preliminary to the reduction of linear covariants we write the formula S as follows:

$$\begin{aligned} S = \{&\phi_0 + b_1\phi_1 + b_1^2\phi_2 + \psi_1(f_1 f_2) + \psi_2(f_1\, Ef_2)\\ &+ \chi_1(Ef_1 f_2) + \chi_2(Ef_1\, Ef_2)\} + \phi_1(b_0 + b_1) + \phi_2(b_0^2 + b_1^2) + \psi_0 a_0\\ &+ \psi_1[a_0 b_0 + (f_1 f_2)] + \psi_2[a_0 b_0^2 + (f_1\, Ef_2)] + \chi_0 a_0^2\\ &+ \chi_1[a_0^2 b_0 + (Ef_1 f_2)] + \chi_2[a_0^2 b_0^2 + (Ef_1\, Ef_2)], \end{aligned}$$

and since there are no linear covariants led by invariants the invariant expression in the brace is congruent to zero.

Therefore the arbitrary linear covariant K is reduced in the form

$$(39)\qquad \begin{aligned} K \equiv\ &\phi_1[f_2] + \phi_2\, E[f_2] + \psi_0 f_1 + \psi_1[f_1 f_2] + \psi_2[f_1\, Ef_2] + \chi_0\, Ef_1\\ &+ \chi_1[Ef_1 f_2] + \chi_2[Ef_1\, Ef_2] \pmod 2. \end{aligned}$$

We have thus completed the proof of the following

THEOREM. *A fundamental system of covariants of the set f_1, f_2 under the modular group G_6* (mod 2) *consists of nineteen quantics, viz.,*

$$f_1,\ f_2,\ Q,\ L,\ Ef_1,\ Ef_2,\ wf_1,\ [f_2],\ [Ef_2],$$

(40) $$[f_1 f_2],\ [f_1 Ef_2],\ [Ef_1 f_2],\ [Ef_1 Ef_2],\ (f_1 f_2),$$

$$q,\ k,\ b_1,\ L',\ Q'.$$

It is an elementary problem to show that the remaining forms involved in the sets (34), (35), (36) are reducible in terms of the quantics (40). Note that it is not stated that the six invariants given form a complete system of pure invariants. The form K' in (37) might involve invariants not reducible in terms of the six.

A theorem concerning modular systems under $G_{(p^2-p)(p^2-1)}$. I proved in previous papers,* under rather strong restrictions as to the point of generality, that a binary form f_m of order m ($m \geqq p^2$) can be expanded covariantively, i.e. given a typical representation in which the denominator function is numerical, in terms of L and Q as argument forms, in such a way that the coefficient forms are first degree formal covariants modulo p of orders $\leqq p^2 - 1$. This expansion was given under the notation

(41) $$f_m = Q^\alpha \phi_1 + Q^{\beta-1} L\chi + \cdots + QL^{\gamma-1}\psi + L^\delta w \pmod p,$$

where ϕ_1 is any one of a number of covariants of a definite type, called principal covariants, each led by a_0. Such an expansion, though not unique, is such that the total number of coefficients involved in the forms $\phi_1, \chi, \cdots$ is at least as great as $m+1$. For brevity write $\phi_1, \chi, \cdots$ under the notation ϕ_i ($i = 1, 2, \cdots$), and let

$$\phi_i = \lambda_{i0} x_1^{\nu_i} + \lambda_{i1} x_1^{\nu_i - 1} x_2 + \cdots + \lambda_{i\nu_i} x_2^{\nu_i}.$$

THEOREM. *A fundamental system modulo* p *of* f_m *is a simultaneous system of the set* ϕ_i ($i = 1, 2, \cdots$).

Suppose that the equation (congruence) (41) is transformed by the general modular substitutions

(42) $$x_1 = \lambda_1 x_1' + \mu_1 x_2', \qquad x_2 = \lambda_2 x_1' + \mu_2 x_2'.$$

Since $L' \equiv (\lambda\mu)^{-1} L$, $Q' \equiv Q \pmod p$ we get

(43) $$f_m' = (a_0', \cdots \mid x_1', x_2')^m \equiv Q^\alpha \phi_1' + Q^{\beta-1} L\phi_2' + \cdots \pmod p,$$

and

$$\phi_i' = \lambda_{i0}' x_1'^{\nu_i} + \lambda_{i1}' x_1'^{\nu_i - 1} x_2' + \cdots + \lambda_{i\nu_i}' x_2'^{\nu_i},$$

where the λ_{ij}' are linear combinations of $\lambda_{i0}, \cdots, \lambda_{i\nu_i}$ and the latter are linear forms in $a_0, \cdots, a_m$. Suppose next that we do not transform (41) as a congruence but that f_m is first transformed by (42) into f_m' and then that the latter is expanded by (41). The result is

*Annals of Mathematics, vol. 19 (1918), p. 201.

$$f'_m = Q^a \phi''_1 + Q^{s-1} L' \chi'' + \cdots + L'^s w'' \pmod p, \tag{44}$$

where

$$\phi''_i = \lambda''_{i0} x_1'^{\nu_i} + \lambda''_{i1} x_1'^{\nu_i - 1} x'_2 + \cdots \qquad (i = 1, 2, \cdots),$$

and λ''_{ij} is the same function of $a'_0, \cdots, a'_m$ that λ_{ij} is of $a_0, \cdots, a_m$; provided (41) satisfies any set of necessary and sufficient conditions whatever which make this expansion unique. Moreover, under these conditions, we get ϵ different sets of induced transformations on the a's and a''s, ϵ being the number of different forms ϕ''_i, viz.,

$$\lambda''_{ij} \equiv \lambda'_{ij} (\lambda\mu)^{a_i} \qquad (i = 1, 2, \cdots, \epsilon). \tag{45}$$

These ϵ sets are together equivalent to the ordinary single set constituting the induced group under (42). Hence, since (45) gives ϵ induced groups for the ϵ forms ϕ_i, and these groups are together equivalent to the transformations on the a's, a concomitant under (42) and (45) i.e. a simultaneous concomitant of the forms ϕ_i $(i = 1, \cdots, \epsilon)$, is a concomitant of f_m and conversely. Hence the theorem is proved.

A complete system of the binary quartic modulo 2. When $m = 4$ (41) is unique without condition and its explicit form is

$$\begin{aligned} f_4 &\equiv Lf_1 + Qf_2 \pmod 2; \\ f_1 &= (a_0 + a_2 + a_3) x_1 + (a_1 + a_2 + a_4) x_2, \\ f_2 &= a_0 x_1^2 + (a_1 + a_2 + a_3) x_1 x_2 + a_4 x_2^2. \end{aligned} \tag{46}$$

Combination of this conclusion with the results numbered (40) gives at once the following

THEOREM. *A fundamental system of formal covariants modulo* 2 *of the binary quartic* f_4 *is composed of nineteen quantics, viz.,* (cf. (46))

$$\begin{aligned} &f_1, f_2, Q, L, Ef_1, Ef_2, wf_1, [f_2], [Ef_2], [f_1 f_2], \\ &[f_1 Ef_2], [Ef_1 f_2], [Ef_1 Ef_2], (f_1 f_2), q, k, (f_4), L', Q'. \end{aligned} \tag{47}$$

In this theory

$$\begin{aligned} q &= a_0^2 + a_0 a_1 + a_0 a_2 + a_0 a_3 + a_0 a_4 + a_1 a_4 + a_2 a_4 + a_3 a_4 + a_4^2, \\ k &= a_0 (a_0 + a_1 + a_2 + a_3 + a_4) a_4, \\ L' &= (a_0 + a_2 + a_3)(a_1 + a_2 + a_4)(a_0 + a_1 + a_3 + a_4), \\ Q' &= a_0 a_2 + a_0 a_4 + a_1 a_2 + a_2 a_4 + a_1 a_3 + a_2 a_3 + a_3 a_4 \\ &\qquad + a_2^2 + a_3^2 + a_4^2 + a_0 a_1 + a_0^2 + a_1^2. \end{aligned} \tag{48}$$

7. Covariants of f_4 modulo 3

The problem of the determination of fundamental covariants led by the seminvariants of the complete system given in § 5 may now be treated.

There exists no quadratic covariant led by an invariant, but $a_0 + a_2$ leads

$$[f_4] = [a_0 + (f_4)]x_1^2 + (a_1 + a_3)x_1 x_2 + [(f_4) + a_4]x_2^2.$$

If we substitute from the induced transformations under $x_1 = x_1' + tx_2'$, $x_2 = x_2'$ in the conjugate to P under the permutations $(a_0 a_4)(a_1 a_3)(a_2)$ and reduce the powers, of the residue t by Fermat's theorem, there results the remarkable skew-covariant

$$[\bar{f}_4] = Px_1^2 + (2a_0 a_2 + a_2 a_4 + a_1^2 - a_3^2)x_1 x_2 + (a_1 a_4 + a_2 a_3 + 2a_3 a_4)x_2^2. \tag{49}$$

From $H + (f_4)^3$ we obtain, by this method,

$$\begin{aligned}[\bar{f}_4] = {} & [H + (f_4)^3]x_1^2 + [a_0 a_1 a_4 + a_2^2 a_3 + 2a_1^2 a_3 + a_1 a_2^2 + 2a_1 a_3^2 \\ & + a_0 a_3 a_4 + a_0 a_2 a_3 + a_1 a_2 a_4]x_1 x_2 \\ & + [(f_4)^3 + a_0 a_4^2 + a_2 a_3^2 - a_1 a_3 a_4]x_2^2.\end{aligned} \tag{50}$$

The algebraical covariant $T_{(3)}$ of order six and degree three with numerical coefficients reduced modulo 3 is a formal covariant whose leader is the seminvariant

$$a_0^2 a_3 + a_0 a_1 a_2 + 2a_1^3.$$

An equivalent covariant which is simpler, in that its leading coefficient is $-\beta$, has the form

$$T = T_{(3)} + f_4[\bar{f}_4],$$

and

$$\begin{aligned}T = {} & -\beta x_1^6 + (a_0 a_2 a_4 + 2a_0^2 a_2 + 2a_0 a_1^2 + 2a_0 a_3^2 + a_0 a_1 a_3 \\ & + 2a_0 a_2^2 + a_0^2 a_4)x_1^5 x_2 + (2a_0 a_2 a_3 + 2a_0 a_3 a_4 + a_1^2 a_3 \\ & + a_1 a_2 a_4 + 2a_1 a_2^2 + 2a_1 a_3^2 + a_1^3)x_1^4 x_2^2 + (a_0 a_3^2 + 2a_1^2 a_4 \\ & + a_0 a_1 a_3 + 2a_1 a_3 a_4 + 2a_0 a_2^2 + a_1^2 a_2 + a_2^2 a_4 + 2a_2 a_3^2)x_1^3 x_2^3 \\ & + (a_0 a_1 a_4 + a_1 a_2 a_4 + 2a_1 a_3^2 + a_2^2 a_3 + 2a_0 a_2 a_3 \\ & + a_1^2 a_3 - a_3^3)x_1^2 x_2^4 + (2a_0 a_4^2 + 2a_1 a_3 a_4 + a_2^2 a_4 + a_3^2 a_4 \\ & + 2a_0 a_2 a_4 + a_2 a_4^2 + a_1^2 a_4)x_1 x_2^5 + (a_3^3 - a_3 a_4^2)x_2^6.\end{aligned} \tag{51}$$

From this one obtains

$$(T) = D + 2\tau \equiv D + 2N + \beta + (a_0 + a_2)P \pmod 3,$$

and

$$
\begin{aligned}
[T] = {} & [-\beta + (T)]x_1^2 + (a_0 a_2^2 + 2a_0 a_1 a_3 + a_0^2 a_4 + 2a_0 a_1^2 \\
& + 2a_0^2 a_2 + a_1 a_3 a_4 + a_1^2 a_2 + 2a_2 a_3^2 + 2a_2^2 a_4 + a_3^2 a_4 \\
& + a_2 a_4^2 + 2a_0 a_4^2) x_1 x_2 + [a_3^3 - a_3 a_4^2 + (T)]x_2^2.
\end{aligned} \tag{52}
$$

Note that $2N + (a_0 + a_2)P + D$, and not β, is the leading coefficient of $[T]$. We find, in addition,

$$
\begin{aligned}
[f_4^2] = {} & [a_0^2 + (f_4^2)]x_1^2 + 2(a_0 a_1 + a_0 a_3 + a_1 a_2 + a_1 a_4 + a_2 a_3 \\
& + a_3 a_4) x_1 x_2 + [(f_4^2) + a_4^2]x_2^2.
\end{aligned} \tag{53}
$$

$$
[Ef_4] = [a_0^3 + (f_4)^3]x_1^2 + (a_1^3 + a_3^3)x_1 x_2 + [(f_4)^3 + a_4^3]x_2^2, \tag{54}
$$

$$
\begin{aligned}
[f_4[\bar{f}_4]] = {} & [a_0 P + 2D + 2\tau]x_1^2 + (2a_0^2 a_2 + 2a_0 a_1^2 + 2a_1^2 a_4 \\
& + 2a_0 a_2^2 + a_2^2 a_4 + a_0 a_3^2 + a_2 a_4^2 \\
& + a_3^2 a_4) x_1 x_2 + [2D + 2\tau + a_1 a_4^2 + a_2 a_3 a_4 + 2a_3 a_4^2]x_2^2.
\end{aligned} \tag{55}
$$

The number of these concomitants which can be constructed by thus superimposing these symbols, one upon another, is infinite. We shall note particularly those given explicitly above, and

$$
\cdot\ [f_4^2[\bar{f}_4]],\ [Ef_4[\bar{f}_4]],\ [f_4 \cdot T],\ \cdots,
$$

whose respective seminvariant leaders are

$$
a_0^2 P + I_1,\ a_0^3 P + I_2,\ -a_0\beta + I_3,\ \cdots,
$$

where I_1, I_2, I_3 are invariants.

It is apparent from the foregoing theory that we have brought a complete system of covariants into view, so to speak, for a covariant led by the arbitrary seminvariant (30) can be derived, part by part, as in the case of covariants (49) to (55), and the parts added, the sum furnishing a reduction similar to that in (37). The formidable calculations required, I have hopes of completing and making the subject of another paper. The main portions of the general problem which remain to be solved are; (1) the determination of the two syzygies described in § 5, needed for the reduction of formula (30) to finite form as regards powers of β and H; (2) determination of all covariants led by invariants which are polynomials in the invariants (30_1); and (3), calculation, from bracket symbols, of the covariants needed to reduce the covariants of the various orders led by the arbitrary seminvariant (30). These can be formed from products of powers of the covariants (49) to (55), combined with f_4. Meanwhile I desire to prove the following theorem, which is of individual importance and is a necessary preliminary to the solution of the complete problem described above.

THEOREM. *A complete set of quadratic covariants led by seminvariants of the general form*

$$(56)\quad \sum_{i=1}^{2}[\,\theta_{0i}+\theta_{1i}\,a_0+\theta_{2i}\,a_0^2+\theta_{3i}\,a_0^3 + (\Delta_{0i}+\Delta_{1i}\,a_0+\Delta_{2i}\,a_0^2+\Delta_{3i}\,a_0^3)\,P\,]\,\beta^{i-1}$$

is composed of

$$(57)\quad \begin{aligned}&[f_4],\ [f_4^2],\ [\,Ef_4],\ [\bar f_4],\ [f_4[\bar f_4]],\ [f_4^2[\bar f_4]],\ [Ef_4[\bar f_4]],\\ &[\,T],\ [f_4\,T],\ [f_4^2\,T],\ [\,Ef_4\,T],\ [\,T[\bar f_4]],\ [f_4\,T[\bar f_4]],\\ &[f_4^2\,T[\bar f_4]],\ [Ef_4\,T[\bar f_4]].\end{aligned}$$

The formula (56) can be written

$$(58)\quad \begin{aligned}&\{\theta_{01}-\theta_{11}(f_4)-\theta_{21}(f_4^2)-\theta_{31}(Ef_4)-\Delta_{11}(2D+2\tau)-\Delta_{21}(f_4^2[\bar f_4])\\ &-\Delta_{31}(Ef_4[\bar f_4])+\theta_{02}(T)+\theta_{12}(f_4\,T)+\theta_{22}(f_4^2\,T)+\theta_{23}(Ef_4\,T)\\ &+\Delta_{02}(T[\bar f_4])+\Delta_{12}(f_4\,T[\bar f_4])+\Delta_{22}(f_4^2\,T[\bar f_4])\\ &+\Delta_{23}(Ef_4\,T[\bar f_4])\}+\theta_{11}(a_0+(f_4))+\theta_{21}(a_0^2+(f_4^2))\\ &+\theta_{31}(a_0^3+(Ef_4))+\Delta_{01}P+\Delta_{11}(a_0P+2D+2\tau)\\ &+\Delta_{21}(a_0^2P+(f_4^2[\bar f_4]))+\Delta_{31}(a_0^3P+(Ef_4[\bar f_4]))\\ &+\theta_{02}(+\beta-(T))+\theta_{12}[\,a_0\beta-(f_4\,T)]+\theta_{22}[\,a_0^2\beta-(f_4^2\,T)]\\ &+\theta_{32}[\,a_0^3\beta-(Ef_4\,T)]+\Delta_{02}[\,\beta P-(T[\bar f_4])]+\Delta_{12}[\,a_0\beta P\\ &-(f_4\,T[\bar f_4])]+\Delta_{22}[\,a_0^2\beta P-(f_4^2\,T[\bar f_4])]\\ &+\Delta_{32}[\,a_0^3\beta P-(Ef_4\,T[\bar f_4])].\end{aligned}$$

Since there are no quadratic covariants led by pure invariants the brace expression in (58) is congruent to zero. The coefficients of the indeterminate invariants θ_{ij}, Δ_{ij} are, in order, respectively, the leading coefficients of the covariants (57), and, by an argument similar to that given in connection with (37), the theorem is therefore proved. Note that the order 2 is the least possible order for a covariant modulo 3 of any form of even order; thus we see that every covariant of this order is rationally and integrally expressible in terms of the fifteen covariants (57) and the six invariants (30_1).

8. ADDENDA CONCERNING INVARIANTS OF VELOCITY AND ACCELERATION IN THE THEORY OF RELATIVITY OF MOTION

The applications of invariant theory noted in this paragraph are grounded upon an algorithm similar to that given in § 1 and may be treated, appro-

priately perhaps, in a paper which gives mention of the general theory and algorithm. For this will not only avoid repetitions, but it will emphasize the relevancy of this research also.

Observations for measurements of velocity, acceleration, time, or distance, made by two observers situated upon two different worlds or systems of reference in space, which are in relative motion with a constant velocity v, give sets of values of these quantities which are connected by linear equations. These equations furnish a means of transferring measurements from one system to another, and three different varieties of transformations of this description are to be considered.

Suppose that the line connecting the two systems S, S' is taken as the axis of abscissas for both systems and that the y- and z-axes upon one system S are respectively parallel to the y'- and z'-axes upon S'. Then, if t and t' are the time variables and c is the velocity of light we can employ the Einstein transformations

$$(59)\quad \Delta: t = \mu(c^2 t' + vx')/c, \qquad x = \mu(vt' + x')c, \qquad y = y', \qquad z = z',$$
$$(\mu = 1/\sqrt{c^2 - v^2}),$$

a known transformation upon the velocities

$$\Omega: \qquad \dot{x}' = (\dot{x} - v) \Big/ \left(1 - \frac{v\dot{x}}{c^2}\right),$$
$$\dot{y}' = \dot{y}k^{-1} \Big/ \left(1 - \frac{v\dot{x}}{c^2}\right),$$
$$\dot{z}' = \dot{z}k^{-1} \Big/ \left(1 - \frac{v\dot{x}}{c^2}\right) \qquad (k = c\mu),$$

where $\dot{x} = \partial x/\partial t$, $\dot{x}' = \partial x'/\partial t'$, etc., and one upon the accelerations given by*

$$A: \quad \ddot{x} = c^{-3}\mu^3(c^2 - v\dot{x})^3\,\ddot{x}',$$
$$(60) \qquad \ddot{y} = c^{-2}\mu^2(c^2 - v\dot{x})^2(\ddot{y}' - v\mu c^{-1}\dot{y}\ddot{x}'),$$
$$\ddot{z} = c^{-2}\mu^2(c^2 - v\dot{x})^2(\ddot{z}' - v\mu c^{-1}\dot{z}\ddot{x}'),$$

in which $\ddot{x} = \partial^2 x/\partial t^2$.

We can reduce the equations Ω to homogeneous form by introducing two fictitious time variables τ, τ' connected with t and t' by functional equations

$$t = f(\tau), \qquad t' = f_1(\tau').$$

For then

$$\frac{\partial x}{\partial t} = \frac{\partial x}{\partial \tau}\frac{\partial \tau}{\partial t}, \qquad \frac{\partial x'}{\partial t'} = \frac{\partial x'}{\partial \tau'}\frac{\partial \tau'}{\partial t'}, \qquad \cdots,$$

* Compare V, p. 48.

and, by taking ρ to be a proportionality factor, we can get the following form,

$$V : \dot{t} = \rho(c^2\dot{t}' + v\dot{x}')/c^2, \qquad \dot{x} = \rho(v\dot{t}' + \dot{x}'),$$

$$\dot{y} = (\mu c)^{-1}\rho\dot{y}', \qquad \dot{z} = (\mu c)^{-1}\rho\dot{z}', \tag{61}$$

$$[\dot{x} = \partial x/\partial\tau, \qquad \dot{x}' = \partial x'/\partial\tau', \qquad \text{etc.}; \qquad dt'/dt = \mu(c^2 - v\partial x/\partial t)/c].$$

*Invariants of V appertaining to an irrational domain.** The components $\dot{y}'$, $\dot{z}'$, are relative covariants. For this reason we are able to treat V as if it were a binary transformation V_1 affecting $\dot{t}$ and $\dot{x}$. Then V_1 has two poles which are the roots of the linear forms

$$\lambda_{\pm 1} = c\dot{t} \pm \dot{x}. \tag{62}$$

Thus $\lambda_{\pm 1}$ are universal covariants of V_1 for an arbitrary velocity v, the invariant relations being

$$\lambda'_{\pm 1} = \left(\frac{c - v}{c + v}\right)^{\mp 1/2}\lambda_{\pm 1}.$$

Consider a binary form f in the variables $\dot{t}$, $\dot{x}$, whose coefficients a_i are homogeneous polynomials in the velocities $\dot{y}$, $\dot{z}$, with coefficients arbitrary functions of the quantities left unaltered by V, these being of the same order in $\dot{y}$, $\dot{z}$,

$$f = a_0\dot{t}^m + ma_1\dot{t}^{m-1}\dot{x} + \cdots + a_m\dot{x}^m.$$

When f is expanded in terms of $\lambda_{\pm 1}$ as argument forms by means of the substitutions

$$\dot{t} = (\lambda_{+1} + \lambda_{-1})/2c, \qquad \dot{x} = c(\lambda_{+1} - \lambda_{-1})/2c,$$

there results a form f' whose coefficients are invariants of a non-absolute type, linear in $a_0, \cdots, a_m$;

$$f' = \sum_{i=0}^{m}\binom{m}{i}\psi_{m-2i}\lambda_{+1}^{m-i}\lambda_{-1}^{i}/(2c)^m. \tag{63}$$

The explicit form of the typical invariant is

$$\binom{m}{i}\psi_{m-2i} = \sum_{j=0}^{m}\sum_{t=0}^{i}(-1)^t\binom{m}{j}\binom{m-j}{i-t}\binom{j}{t}a_j c^j \quad (i = 0, \cdots, m), \tag{64}$$

the invariant relations being

$$\psi'_{m-2i} = \left(\frac{c + v}{c - v}\right)^{(m-2i)/2}\psi_{m-2i}.$$

A general theory of the invariants derived in this way,† which the author has developed, shows that systems of absolute concomitants of f can be

* Cf. these Transactions, vol. 18 (1917), p. 443, and vol. 20 (1919), p. 203.

† Annals of Mathematics, vol. 20 (1918), p. 125.

derived by forming properly chosen products of powers of the invariants ψ_{m-2i}; that is, systems which are absolute except for a power of the proportionality factor ρ occurring in the invariant relations.

Concomitants belonging to a rational domain. The general form of such a product is

$$I_r = \phi_{m-2i_1}^{r_{i_1}} \phi_{m-2i_2}^{r_{i_2}} \cdots \lambda_{+1}^{s_1} \lambda_{-1}^{s_2},$$

and the following product is conjugate to I_r

$$I_{-r} = \phi_{-(m-2i_1)}^{r_{i_1}} \phi_{-(m-2i_2)}^{r_{i_2}} \cdots \lambda_{+1}^{s_2} \lambda_{-1}^{s_1}.$$

According to the theory $I_r \pm I_{-r}$ is an absolute concomitant provided the positive integral exponents satisfy the linear diophantine equation

$$\begin{aligned} &s_2 + r_0 m + r_1(m-2) + r_2(m-4) + \cdots \\ &\qquad = \cdots r_{m-2}(m-4) + r_{m-1}(m-2) + r_m m + s_1 \end{aligned} \tag{65}$$

and a complete system is furnished by the finite complete set of irreducible solutions of (65).

In the case of a binary quartic ($m = 4$) I have found that the number of irreducible solutions is 12, as shown in the table below.

	r_0	r_1	r_2	r_3	r_4	s_1	s_2
ν	0	1	0	1	0	0	0
ξ	1	0	0	0	1	0	0
o	0	0	0	0	0	1	1
π	0	0	1	0	0	0	0
$\rho_{\pm 1}$	0	1	0	0	0	2	0
	0	0	0	1	0	0	2
$\sigma_{\pm 1}$	1	0	0	2	0	0	0
	0	2	0	0	1	0	0
$\tau_{\pm 1}$	1	0	0	1	0	2	0
	0	1	0	0	1	0	2
$v_{\pm 1}$	1	0	0	0	0	4	0
	0	0	0	0	1	0	4

The fundamental system is

$$\begin{aligned} &\nu = \psi_2 \psi_{-2}, \qquad \xi = \psi_4 \psi_{-4}, \qquad o = c^2 \left(\frac{\partial t}{\partial \tau}\right)^2 - \left(\frac{\partial x}{\partial \tau}\right)^2 = \lambda_{+1} \lambda_{-1}, \\ &\pi = \psi_0, \qquad \rho_{\pm 1} = \psi_2 \lambda_{+1}^2 \pm \psi_{-2} \lambda_{-1}^2, \\ &\sigma_{\pm 1} = \psi_4 \psi_{-2}^2 \pm \psi_{-4} \psi_2^2, \qquad \tau_{\pm 1} = \psi_4 \psi_{-2} \lambda_{+1}^2 \pm \psi_{-4} \psi_2 \lambda_{-1}^2, \\ &v_{\pm 1} = \psi_4 \lambda_{+1}^4 \pm \psi_{-4} \lambda_{-1}^4, \end{aligned} \tag{66}$$

and is given explicitly by the following functions:

$$\psi_0 = a_0 - 2a_2 c^2 + a_4 c^4,$$

$$\psi_{\pm 2} = a_0 \pm 2a_1 c \mp 2a_3 c^3 - a_4 c^4,$$

$$\psi_{\pm 4} = a_0 \pm 4a_1 c + 6a_2 c^2 \pm 4a_3 c^3 + a_4 c^4.$$

Invariants of acceleration

Two of the three poles of the transformations (60) are coincident. For this reason we cannot derive three linearly independent linear universal covariants of this transformation, but only two, and thus a ternary form in the component accelerations $\ddot{x}, \ddot{y}, \ddot{z}$, cannot be expanded covariantively in the form analogous to f' above, and no analogue of the covariant theory of velocities exists, in the ternary realm. The two linear universal covariants are

$$\ddot{x}, \quad \dot{z}\ddot{y} - \dot{y}\ddot{z},$$

and the second one can be employed to reduce the transformation (60) to a canonical form in which both quantities $\ddot{x}_1, \ddot{y}_1$ are relative covariants. The invariant problem then becomes binary.

Let

$$\dot{z}\ddot{y} - \dot{y}\ddot{z} = \ddot{y}_1,$$

this notation being adopted for uniformity. Then $\ddot{y}_1 = \mu^2 \delta^2 \ddot{y}_1'$, where $c\delta = c^2 - v\dot{x}$. Accordingly the transformation (60) is equivalent to the following:

$$\theta : \begin{cases} \ddot{x} = \mu^3 \delta^3 \ddot{x}', \\ \ddot{y}_1 = \mu^2 \delta^2 \ddot{y}_1', \\ \ddot{z} = \mu^2 \delta^2 (- v\mu c^{-1} \dot{z}\ddot{x}' + \ddot{z}'). \end{cases} \tag{67}$$

The substitutions on $\ddot{x}, \ddot{z}$, form a binary transformation θ_1 and the linear universal covariants of θ_1 are

$$\begin{aligned} f_{+1} &= 2\mu^2 \delta^2 [- \mu v c^{-1} \dot{z}\ddot{x} + (1 - \mu\delta)\ddot{z}] \equiv C\ddot{x} + D\ddot{z}, \\ f_{-1} &= - 2\mu^3 v c^{-1} \delta^2 \dot{z}\ddot{x} \equiv C\ddot{x}. \end{aligned} \tag{68}$$

Then, if $f_m = (a_0, \cdots, a_m)(\ddot{x}, \ddot{z})^m$ is a binary quantic with constant coefficients, the expansion of f_m in terms of f_{+1}, f_{-1} as arguments has for coefficients linear invariants of f_m under θ_1, the totality of $m + 1$ furnishing a complete system of the non-absolute type.

In a paper published in 1917 I gave the invariants of this type* for the general transformation

$$x_1 = \alpha_1 x_1' + \alpha_2 x_2',$$

$$x_2 = \beta_0 x_1' + \beta_1 x_2',$$

* These Transactions, vol. 18 (1917), p. 443.

and the orders $m = 1, 2, 3$. From these we get the invariants of θ_1 by the particularizing substitutions

$$\alpha_1 = \mu^3 \delta^3, \qquad \alpha_2 = 0, \qquad \beta_0 = -\mu^3 \delta^2 c^{-1} v\dot{z}, \qquad \beta_1 = \mu^2 \delta^2. \tag{69}$$

The sets for the orders $m = 1$, $m = 2$, for example are as follows:

$$m = 1$$

$$\begin{aligned} \phi_1' &= a_1, \\ \phi_{-1}' &= (1 - \mu\delta)\, a_0 + \mu v c^{-1} \dot{z} a_1. \end{aligned} \tag{70}$$

$$m = 2$$

$$\begin{aligned} \phi_2' &= a_2, \\ \phi_0' &= (1 - \mu\delta)\, a_1 + \mu v c^{-1} \dot{z} a_2, \\ \phi_{-2}' &= (1 - \mu\delta)^2 a_0 + 2\mu v c^{-1} \dot{z} (1 - \mu\delta)\, a_1 + \mu^2 v^2 c^{-2} \dot{z}^2 a_2. \end{aligned}$$

For these invariants of f_m under θ_1 we find the following general formula:

$$\phi_{m-2r}' = \sum_{s=0}^{r} \binom{r}{s} (1 - \mu\delta)^{r-s} (\mu v c^{-1} \dot{z})^s a_{m-r+s} \qquad (r = 0, \cdots, m). \tag{71}$$

We also give the corresponding invariant relations, viz.,

$$f_{+1}' = (\mu^2 \delta^2)^{-1} f_{+1}, \qquad f_{-1}' = (\mu^3 \delta^3)^{-1} f_{-1}, \tag{72}$$

$$(\phi_{m-2r}')' = (\mu \cdot \delta)^{2m+r} \phi_{m-2r}'. \tag{73}$$

UNIVERSITY OF PENNSYLVANIA

PROPERTIES OF THE SUBGROUPS OF AN ABELIAN PRIME POWER GROUP WHICH ARE CONJUGATE UNDER ITS GROUP OF ISOMORPHISMS*

BY

G. A. MILLER

1. INTRODUCTION

Two operators or two subgroups of a group will be called I-conjugate whenever they are conjugate under the group of isomorphisms of this group. It is well known that all the cyclic subgroups of highest order contained in any abelian group are I-conjugate. A necessary and sufficient condition that all the cyclic subgroups of any other order contained in an abelian group G of order p^m, p being a prime number, are I-conjugate is that all the invariants of G are equal to each other. When this condition is satisfied we shall prove, in particular, that the number of the subgroups of order p^α, $\alpha \leqq m$, contained in G is equal to the number of its subgroups of index p^α.

It is known that the ϕ-subgroup of G is composed of the pth power of every operator of G. The concept of ϕ-subgroup in connection with abelian groups can readily be extended by considering the characteristic subgroup of G composed of the p^αth power of every operator of G. This subgroup may be called the ϕ_α-subgroup of G. It is composed of all the operators of G which are found in each one of its subgroups of index p^α, and all the operators of highest order in a ϕ_α-subgroup of G are I-conjugate. In particular, the ϕ_α-subgroup generated by operators of order p is the fundamental characteristic subgroup of G.† In fact, when $p > 2$ the ϕ_α-subgroup which is generated by operators of order p^r is contained in every characteristic subgroup of G which involves operators of order p^r.

The quotient group of G corresponding to a ϕ_α-subgroup may be called a ϕ_α-quotient group. This quotient group is simply isomorphic with the subgroup of G generated by all its operators whose orders divide p^α. Hence the following:

THEOREM. *The number of the subgroups of index p^α contained in any abelian*

* Presented to the Society, December 31, 1919.

† G. A. Miller, American Journal of Mathematics, vol. 27 (1905), p. 16.

group of order p^m is equal to the number of the subgroups of index p^a contained in its subgroup composed of all its operators whose orders divide p^a.

As a special case of this theorem it may be noted that the number of subgroups of index p contained in any abelian group is equal to the number of its subgroups of order p, since it is known that the number of the subgroups of index p contained in the abelian group of order p^m and of type $(1, 1, 1, \cdots)$ is equal to the number of its subgroups of order p. When $p^a \leqq$ the smallest invariant of G it results from this theorem and the theorem proved in the following section that the number of subgroups of order p^a contained in G is equal to the number of its subgroups of index p^a.

Among the theorems of section 3 we may here note the one which establishes the fact that if h is the order of the largest cyclic subgroup of any abelian group and if d is a divisor of h then the number of sets of I-conjugate operators of order d contained in this group is the same as the number of its sets of I-conjugate operators of order h/d. In the special case when $d = 1$ this theorem reduces to the one according to which the operators of highest order in any abelian group are I-conjugate.

If the ϕ_a-subgroup of G is of type $(\alpha_1', \alpha_2', \cdots, \alpha_\lambda')$ and if the ϕ_a-quotient group is of type $(\alpha_1'', \alpha_2'', \cdots, \alpha_\lambda'')$; where some of these α's may be 0 then these α's may be so arranged that

$$\alpha_1' + \alpha_1'' = \alpha_1, \qquad \alpha_2' + \alpha_2'' = \alpha_2, \qquad \cdots, \qquad \alpha_\lambda' + \alpha_\lambda'' = \alpha_\lambda$$

where $(\alpha_1, \alpha_2, \cdots, \alpha_\lambda)$ represents the type of G. Two such subgroups of G may be said to be of *complementary types*. It should be noted that in general the type of a subgroup of an abelian group of order p^m does not determine completely the type of its complementary subgroup.

2. Abelian groups of order p^m all of whose invariants are equal to each other

When all the invariants of G are equal to each other all the operators of the same order contained in G are I-conjugate, and hence it results that the number of the characteristic subgroups of G is one less than the number of the different orders of the operators of G. A necessary and sufficient condition that two subgroups of G are I-conjugate is that they are of the same type, and any subgroup and the corresponding quotient group are of complementary types since every operator of such a G is a power of some operator of highest order contained in G.

The theorem noted in the Introduction as regards the equality between the number of subgroups of a given order and the number of the subgroups of the same index is clearly a corollary of the theorem that the number of the subgroups of a given type is the same as the number of the subgroups of the

complementary type. To simplify a proof of this theorem we shall first consider the special case when all the invariants of G are equal to p.

Since the number of the subgroups of order p is known to be equal to the number of the subgroups of index p, and the type of one of the former subgroups is complementary to the type of one of the latter, the theorem in question will be proved for this special type of groups provided it is proved that the number of the subgroups of order p^α is equal to the number of the subgroups of index p^α whenever the number of the subgroups of order $p^{\alpha-1}$ is equal to the number of the subgroups of index $p^{\alpha-1}$, $m > \alpha > 1$.

As G is supposed to be of type $(1, 1, 1, \cdots)$ the number of its subgroups of order p^α which contain the same subgroup of order $p^{\alpha-1}$ is equal to the number of subgroups of order p in the abelian group of order $p^{m-\alpha+1}$ and of type $(1, 1, 1, \cdots)$. On the other hand, the number of the subgroups of index p^α which are contained in the same subgroup of index $p^{\alpha-1}$ is equal to the number of subgroups of index p in the abelian group of order $p^{m-\alpha+1}$ and of type $(1, 1, 1, \cdots)$. As this group is known to have exactly as many subgroups of order p as of index p it results that if we count each subgroup of order p^α as many times as it contains subgroups of order $p^{\alpha-1}$ we obtain the same sum as if we count each subgroup of index p^α as many times as it appears in some subgroup of index $p^{\alpha-1}$ since it was assumed that the number of subgroups of order $p^{\alpha-1}$ contained in G is equal to the number of its subgroups of index $p^{\alpha-1}$.

By the method of counting noted in the preceding paragraph each subgroup of order p^α is counted $(p^\alpha - 1)/(p - 1)$ times, and each subgroup of index p^α is counted the same number of times. Hence it results that the number of the subgroups of order p^α contained in G is equal to the number of the subgroups of order $p^{\alpha-1}$ multiplied by

$$\frac{p^{m-\alpha+1} - 1}{p^\alpha - 1}.$$

The number of the subgroups of index p^α may be found by multiplying the number of the subgroups of index $p^{\alpha-1}$ by the same factor. This completes the proof by induction of the theorem that in an abelian group of order p^m and of type $(1, 1, 1, \cdots)$ the number of the subgroups of a given type is equal to the number of the subgroups of the complementary type.

Having proved the theorem in question when G is of type $(1, 1, 1, \cdots)$ we proceed to give a proof of it when all the invariants of G are equal to p^{α_1}, $\alpha_1 > 1$. In this case the subgroups of order p^α contained in G can be of different types, but it may be assumed that the number of the subgroups of order $p^{\alpha-1}$ and of any possible type is equal to the number of the subgroups of index $p^{\alpha-1}$ and of the complementary type.

To find the number of the subgroups of order p^a and of a given type we may count each of these subgroups as many times as it contains subgroups of index p such that the orders of their independent generators differ from the orders of the independent generators of one of the former subgroups H only as regards the fact that one of the largest independent generators of H is replaced by one whose order is equal to the order of this generator divided by p. It should be noted that not all the subgroups of index p under H are thus counted but only those which are of the particular type just stated.

The number of these subgroups contained in H is $(p^\lambda - 1)/(p - 1)$, where λ is the number of the largest invariants of H, since all of these subgroups of index p contain the characteristic subgroup of H generated by all its operators which are not of highest order. This is also the number of the subgroups which contain a given subgroup of index p^a whose type is complementary to that of H and are of a type which is complementary to the type of one of the given subgroups of order p^{a-1}. Hence it results that if we count each subgroup which is of the same type as H as many times as it contains a subgroup of order p^{a-1} and of the given type we obtain the total number of subgroups which are I-conjugate with H multiplied by a given number k, and if we count each subgroup which is of a type complementary to that of H as many times as it is found in a subgroup of index p^{a-1} which is of a type complementary to the given subgroups of order p^{a-1} we obtain k times the total number of the subgroups whose type is complementary to that of H.

The number of the subgroups of G which have the same type as H and which contain a particular subgroup of order p^{a-1} and of the given type is $(p^{\gamma-\lambda+1} - p^{\gamma-\lambda'})/(p - 1)$, λ' being the number of the independent generators of H whose order is equal to or greater than the order of the largest independent generator of H divided by p and γ being the number of the independent generators of G. This results from the facts that all such subgroups are found in the characteristic subgroup of G which is generated by all the operators of G whose orders do not exceed the order of the largest operator of H and that the quotient group of this subgroup with respect to one of the given subgroups of order p^{a-1} contains $\gamma - \lambda + 1$ independent generators. The subgroups of order p in this quotient group which correspond to groups having only $\lambda - 1$ invariant which are equal to the largest invariant of H generate a group of order $p^{\gamma-\lambda'}$.

In a similar way it can be proved that the number of the subgroups whose type is complementary to that of H and which are found in a given subgroup of index p^{a-1} which is complementary to one of the given subgroups of order p^{a-1} is also $(p^{\gamma-\lambda+1} - p^{\gamma-\lambda'})/(p - 1)$. This can also be proved by finding the number of ways in which the independent generators of the former subgroup can be selected from the operators of the latter and by dividing this

number by the number of ways in which these independent generators can be selected from the operators of the former subgroup.

From what precedes it follows that if the number of the subgroups of order p^{a-1} and of the given type is equal to the number of the subgroups of complementary type then if we count each subgroup of G which is simply isomorphic with H as many times as it contains a subgroup of order p^{a-1} and of the given type we get the same sum as when we count each of the subgroups of G whose type is complementary to the type of H as many times as it appears in a subgroup of index p^{a-1} whose type is complementary to that of the given subgroups of order p^{a-1}. As by this process each of these two types of subgroups is counted the same number of times we have established the following:

THEOREM. *If all the invariants of an abelian group G of order p^m are equal to each other then the number of the subgroups of a given type contained in G is equal to the number of the subgroups which are of the complementary type.*

As an important corollary of this theorem it may be noted that *the number of the distinct subgroups of order p^a contained in any abelian group of order p^m whose invariants are equal to each other is the same as the number of its distinct subgroups of index p^a.*

3. NUMBER OF SETS OF I-CONJUGATE OPERATORS OF G

Let $p^{a_1}, p^{a_2}, \cdots, p^{a_\lambda}$ represent all the different invariants of G arranged in descending order of magnitude. There may be more than one invariant of G which is equal to any one of these numbers but no two of these numbers are supposed to be equal to each other. Let $I_{a_1-\beta}$, $\alpha_1 > \beta \geqq 0$, represent the substitution group according to which the operators of order $p^{a_1-\beta}$ contained in G are permuted under I. It is known that I_{a_1} is transitive and that I_1 contains exactly λ transitive constituents. It was noted in the Introduction that one of the transitive constituents of $I_{a_1-\beta}$ corresponds to the ϕ_β-subgroup of G.

To simplify the proof of the theorem that there are just as many transitive constituents in $I_{a_1-\beta}$ as in I_β we shall first prove the special case of this theorem which relates to I_1 and I_{a_1-1}. It was noted above that I_1 contains λ transitive constituents and that the operators of highest order in the ϕ_1-subgroup of G correspond to a single transitive constituent of I_1. If we adjoin to the ϕ_1-subgroup of G all the smallest independent generators of G there results a group in which the operators of highest order correspond to two of the transitive constituents of I_1. The group thus obtained can be extended by the operators in a set of independent generators of G which are next to the lowest order and the resulting group will contain three sets of I-conjugate operators of order p^{a_1-1} whenever $\lambda > 2$, etc. Finally, the operators of highest order in the group generated by all the operators which are not of highest order in a set

of independent generators of G, together with the pth powers of the operators of highest order in this set correspond to λ transitive constituents of I_1. This completes a proof of the special theorem that I_1 and I_{α_1-1} have the same number λ of transitive constituents.

Before proving the general theorem under consideration it may be desirable to indicate how the total number of the sets of I-conjugate operators of a given order p^r can be determined. This is evidently equivalent to the determination of the number of the transitive constituents of I_r and also of the number of the characteristic subgroups of G which are generated by operators of order p^r. The smallest of these transitive constituents of I_r, when, $p > 2$, corresponds to all the operators of order p^r contained in G which are powers of operators of order p^{α_1}. To the characteristic subgroup K_1 generated by these operators we adjoin a characteristic subgroup of G which satisfies the following two conditions: Its operators of largest order have the same order as the operators of the smallest order in K_1, and its order is as small as possible so as to contain operators not found in K_1. We thus obtain a characteristic subgroup whose operators correspond to two transitive constituents of I_r.

If the operators of lowest order in this extended characteristic subgroup do not include all the operators of the same order contained in G this process is repeated. Each repetition increases by one the number of I-conjugate operators in the characteristic subgroup thus extended. After the operators of lowest order in such a group include all the operators of the same order appearing in G we adjoin to K_1 the smallest characteristic subgroup of G generated by operators of next to the lowest order found in K_1 and involving operators not found in K_1. The characteristic subgroup K_2 thus constructed is first treated just like K_1 was treated to obtain successive characteristic subgroups, each involving one more set of I-conjugate operators of highest order than the preceding one. We then build K_3 on K_2 just as K_2 was obtained by extending K_1, etc.

This general method may be illustrated by finding the number of the I-conjugate operators of order p^3 in the abelian group of order p^{21} and of type $(6, 5, 4, 3, 2, 1)$. The subgroup K_1 is of order p^6 and may be extended three times, in order, by operators of order p. We thus obtain 4 sets of I-conjugate operators of order p^3. The group K_2 is of order p^8 and can be extended twice by operators of order p. We thus obtain 3 additional sets of I-conjugate operators of order p^3. The group K_3 is of order p^{10} and can be extended once by operators of order p while K_4 is of order p^{12}. As these groups contain together 3 additional sets of I-conjugate operators of order p^3 it results that in the characteristic subgroups which involve separately only one invariant which is equal to p^3 there are 10 sets of I-conjugate operators of order p^3.

When the characteristic subgroups which are generated by operators of order p^3 and contain two invariants which are equal to p^3 are considered we evidently have a repetition of a part of the preceding problem, and the number of additional sets of I-conjugate operators of order p^3 is 6. When the characteristic subgroups which are generated by operators of order p^3 involve three invariants which are equal to p^3 there are 3 additional sets of such I-conjugate operators and there is one additional set when there are four invariants which are equal to p^3. Hence the total number of transitive constituents in I_3 for the given group is $10 + 6 + 3 + 1 = 20$. This is also equal to the number of the characteristic subgroups which are generated by operators of order p^3.

Having given and illustrated a method by which the number of the sets of I-conjugate operators of order p^r can be found we proceed to prove that the number of the sets of I-conjugate operators of order p^{α_1-r} is exactly the same. For this purpose it may evidently be assumed that $\alpha_1 \geqq 2r$. The proof is practically the same as the one given above. In the present case we begin with the subgroup generated by the p^r powers of the operators in a set of independent generators of G_1 and note that the operators of order p^{α_1-r} contained in this subgroup constitute a single set of I-conjugate operators of G since they are separately generated by operators of order p^{α_1}. To obtain a subgroup which contains just two sets of I-conjugate operators of order p^{α_1-r} we adjoin to this subgroup the p^{r-1} powers of the smallest independent generators of G, provided their order is at least equal to p^r. If their order is less than p^r we adjoin the smallest characteristic subgroup of G which is generated by operators of order p but is not found in the subgroup already constructed. This subgroup can be extended so as to obtain a subgroup containing one more set of I-conjugate operators just as in the preceding case, and hence it is clear that the number of sets of I-conjugate operators of order p^{α_1-r} is exactly the same as the number of such sets of operators of order p^r.

In an abelian group whose order is not the power of a prime number the number of the sets of I-conjugate operators of a given order which is represented by operators of the group is equal to the product of the numbers of such sets for the prime power constituents of this order. Hence the method illustrated above enables one to determine directly the number of the different sets of I-conjugate operators of every order represented by the operators of the group. In particular, it results that if the largest order h of an operator of any abelian group is divisible by d the number of the different sets of its I-conjugate operators of order d is equal to the number of its different sets of I-conjugate operators of order h/d, as was noted in the Introduction.

It may be added that the present article extends along various lines the

results obtained in an article by the present writer, published in volume 27 (1905), of the American Journal of Mathematics, page 15, under the title "Determination of all the characteristic subgroups of an abelian group." Each set of I-conjugate operators of a group generates a characteristic subgroup and every characteristic subgroup is composed of one or more than one complete set of I-conjugate operators. Hence the problem of determining the number of the characteristic subgroups of an abelian group is equivalent to the determination of the number of the transitive constituents in the group of isomorphisms of this abelian group.

ON THE ORDER OF MAGNITUDE OF THE COEFFICIENTS IN TRIGONOMETRIC INTERPOLATION*

BY

DUNHAM JACKSON

1. **Introduction.** It is well known that there is a remarkable and far-reaching parallelism between the properties of the trigonometric interpolating formula determined by the values of a given function at a set of equally spaced points, and those of the Fourier's series for the same function.† This parallelism is closer than the correspondence of sums and integrals in the coefficients would indicate at first sight. Let $f(x)$ be the given function, assumed to have the period 2π. Let the Fourier's series, to terms of the pth order, be denoted by

$$s_p(x) = \frac{a_0}{2} + \sum_{n=1}^{p} (a_n \cos nx + b_n \sin nx).$$

Its coefficients are

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x) \cos nx\, dx, \qquad b_n = \frac{1}{\pi}\int_0^{2\pi} f(x) \sin nx\, dx.$$

Let the finite trigonometric sum of order p which agrees in value with $f(x)$ at $2p + 1$ points equally spaced over the interval from 0 to 2π be

$$\sigma_p(x) = \frac{a_{0p}}{2} + \sum_{n=1}^{p} (a_{np} \cos nx + b_{np} \sin nx);$$

in particular, it will be assumed that the points of coincidence are the points $x_k = 2k\pi/(2p + 1)$, $k = 0, 1, \cdots, 2p$. The coefficients here are the sums

$$(1) \quad a_{np} = \frac{2}{2p+1}\sum_{k=0}^{2p} f(x_k) \cos nx_k, \qquad b_{np} = \frac{2}{2p+1}\sum_{k=0}^{2p} f(x_k) \sin nx_k.$$

The resemblance of these to the integrals for a_n and b_n is apparent; and it is

* Presented to the Society, December 30, 1919.

† Cf., e.g., Bôcher, *Introduction to the theory of Fourier's series*, Annals of Mathematics, vol. 7 (1905–06), pp. 81–152; pp. 132–137; Faber, *Ueber stetige Funktionen*, Mathematische Annalen, vol. 69 (1910), pp. 372–443; pp. 417–443; D. Jackson, *On the accuracy of trigonometric interpolation*, these Transactions, vol. 14 (1913), pp. 453–461.

seen that if n is held fast and p increases without limit,

$$\lim_{n=\infty} a_{np} = a_n, \qquad \lim_{n=\infty} b_{np} = b_n.$$

For any fixed value of p, however, if n is nearly equal to p, the variations of $\cos nx$ and $\sin nx$ are only very imperfectly represented by the values of these functions at a set of $2p+1$ points, so that the similarity between s_p and σ_p has to be regarded as a consequence of the special properties of the trigonometric functions, not merely of the definition of an integral as the limit of a sum.

The present paper is concerned with the order of magnitude of the coefficients in the interpolating formula, in case the function $f(x)$ is of limited variation, or has a derivative of limited variation. The method of partial summation is used to give results analogous to those which Picard* obtains for the Fourier's series by the second law of the mean and integration by parts. Incidentally, in the case of a derivative of limited variation, a relatively simple proof is found for the convergence of the Fourier's series, as well as of the interpolating function, to the value† $f(x)$.

2. **Order of magnitude of the coefficients for a function of limited variation.** Let $f(x)$ be a function of period 2π, having limited variation in any finite interval, its total variation in the interval from 0 to 2π being V. Then,‡ if a_n and b_n are the coefficients in the Fourier's series for $f(x)$, as above,

$$|a_n| \leqq \frac{K_1}{n}, \qquad |b_n| \leqq \frac{K_1}{n}, \qquad n > 0,$$

where K_1 is independent of n, and may be taken equal to V/π. This can be seen as follows. The function $f(x) - f(0)$ can be expressed in the form

$$f(x) - f(0) = f_1(x) - f_2(x),$$

where $f_1(x)$ and $f_2(x)$, the positive and negative variations of $f(x)$ from 0 to x, are non-decreasing functions which vanish for $x = 0$. Since

$$f(2\pi) = f(0) \qquad \text{and} \qquad f_1(2\pi) + f_2(2\pi) = V,$$

it follows that

$$f_1(2\pi) = f_2(2\pi) = \frac{V}{2}.$$

The value of a_n is

$$a_n = \frac{1}{\pi}\left[\int_0^{2\pi} f_1(x)\cos nx\,dx - \int_0^{2\pi} f_2(x)\cos nx\,dx + f(0)\int_0^{2\pi}\cos nx\,dx\right].$$

* *Traité d'analyse*, pp. 252–253, 255–256 (second edition).

† It is immediately apparent from the order of magnitude of the coefficients that the Fourier's series is uniformly convergent, but the proof that the sum of the series is identical with $f(x)$ usually requires more attention.

‡ Cf. Picard, loc. cit., pp. 252–253.

The third integral in the brackets is zero, for $n > 0$. By the second law of the mean,

$$\int_0^{2\pi} f_1(x) \cos nx\, dx = f_1(2\pi) \int_\xi^{2\pi} \cos nx\, dx = \frac{V}{2}\left[\frac{\sin nx}{n}\right]_\xi^{2\pi} = -\frac{V}{2} \cdot \frac{\sin n\xi}{n},$$

where $0 \leqq \xi \leqq 2\pi$, and hence

$$\left| \int_0^{2\pi} f_1(x) \cos nx\, dx \right| \leqq \frac{V}{2n}.$$

Similarly,

$$\left| \int_0^{2\pi} f_2(x) \cos nx\, dx \right| \leqq \frac{V}{2n},$$

and therefore

$$|a_n| \leqq \frac{V}{n\pi}.$$

A corresponding computation for b_n would seem at first sight to give only the inequality $|b_n| \leqq 2V/(n\pi)$, since the value of $\int \sin nx\, dx$, extended from an unknown lower limit to 2π, might be as great as 2. For any particular value of n, however, the integral defining b_n can equally well be extended over an interval of length 2π beginning at the point $\pi/(2n)$, and so terminating at a point where $\cos nx$ vanishes; and by expressing $f(x) - f(\pi/(2n))$ as the difference of two non-decreasing functions which vanish for $x = \pi/(2n)$, it can be shown that

$$|b_n| \leqq \frac{V}{n\pi}.$$

An analogous result for the coefficients a_{np}, b_{np} is as follows:

THEOREM I. *If $f(x)$ is a function of period 2π which is of limited variation in any finite interval, then*

$$|a_{np}| \leqq \frac{L_1}{n}, \qquad |b_{np}| \leqq \frac{L_1}{n}, \qquad n > 0,$$

where L_1 is independent of n and p.

If V is the total variation of $f(x)$ in the interval from 0 to 2π, the constant L_1 can be taken equal to V.

It will be convenient to consider the sums

$$\sum f(x_k) \cos nx_k \qquad \text{and} \qquad \sum f(x_k) \sin nx_k$$

of (1) together, as the real part and the coefficient of the pure imaginary part of the expression

$$\Sigma_0 = \sum_{k=0}^{2p} f(x_k)\, e^{nx_k i}.$$

Let

$$w_q = \sum_{k=0}^{q} e^{nx_k i} = \sum_{k=0}^{q} e^{k\cdot[2n\pi i/(2p+1)]} \qquad (q = 0, 1, \cdots, 2p);$$

then

$$\begin{aligned}\Sigma_0 &= f(x_0)\cdot w_0 + f(x_1)(w_1 - w_0) + f(x_2)(w_2 - w_1) \\ &\qquad + \cdots + f(x_{2p})(w_{2p} - w_{2p-1}) \\ &= [f(x_0) - f(x_1)]w_0 + [f(x_1) - f(x_2)]w_1 \\ &\qquad + \cdots + [f(x_{2p-1}) - f(x_{2p})]w_{2p-1} + f(x_{2p})\cdot w_{2p}. \end{aligned} \tag{2}$$

The last term written down vanishes, because* $w_{2p} = 0$. Let H be a number such that

$$|w_q| \leqq H$$

for $q = 0, 1, \cdots, 2p-1$. Then

$$\begin{aligned}|\Sigma_0| \leqq [&|f(x_0) - f(x_1)| + |f(x_1) - f(x_2)| \\ &+ \cdots + |f(x_{2p-1}) - f(x_{2p})|]H,\end{aligned}$$

or, as the sum in brackets can not exceed V,

$$|\Sigma_0| \leqq VH. \tag{3}$$

It remains to form an estimate of the magnitude of H.

By applying to w_q the formula for the sum of $q+1$ terms of a geometric progression, it is seen that

$$w_q = \frac{1 - e^{[2(q+1)n\pi i/(2p+1)]}}{1 - e^{[2n\pi i/(2p+1)]}}. \tag{4}$$

The second term in the numerator is represented by a point on the unit circle in the complex plane, so that the absolute value of the numerator can not exceed 2. It is evident geometrically (and of course can readily be proved analytically) that the absolute value of the denominator is $2\sin[n\pi/(2p+1)]$. Since $n/(2p+1) < \frac{1}{2}$, and since $\sin y/y$ decreases steadily as y increases from 0 to $\pi/2$,

$$\sin\frac{n\pi}{2p+1} = \frac{n\pi}{2p+1}\cdot\frac{\sin[n\pi/(2p+1)]}{n\pi/(2p+1)} > \frac{n\pi}{2p+1}\cdot\frac{\sin(\pi/2)}{\pi/2},$$

that is,

$$\sin\frac{n\pi}{2p+1} > \frac{2n}{2p+1}.$$

* This is a consequence of well-known properties of the sine and cosine, and, in particular, can be inferred immediately from the formula (4) below.

It is found consequently that

$$|w_q| \leqq \frac{2p+1}{2n},$$

and the quantity on the right can be substituted for H in (3).

Finally, since a_{np} and b_{np} are respectively the real part and the coefficient of the pure imaginary part of $2\Sigma_0/(2p+1)$,

$$|a_{np}| \leqq \frac{V}{n}, \qquad |b_{np}| \leqq \frac{V}{n}.$$

3. **Order of magnitude of the coefficients for a function having a first derivative of limited variation.** Let it be assumed now that $f(x)$ has a first derivative which is of limited variation in any finite interval, and let V in this case stand for the total variation of the derivative in the interval from 0 to 2π. By integration by parts,*

$$\int_0^{2\pi} f(x) \cos nx\, dx = -\frac{1}{n}\int_0^{2\pi} f'(x) \sin nx\, dx,$$

$$\int_0^{2\pi} f(x) \sin nx\, dx = \frac{1}{n}\int_0^{2\pi} f'(x) \cos nx\, dx,$$

the additional terms which would otherwise appear on the right vanishing because of the periodicity of the functions involved. Since $f'(x)$ satisfies the hypotheses that were imposed on $f(x)$ at the beginning of the preceding section,

$$\left|\int_0^{2\pi} f'(x) \sin nx\, dx\right| \leqq \frac{V}{n}, \qquad \left|\int_0^{2\pi} f'(x) \cos nx\, dx\right| \leqq \frac{V}{n},$$

and

$$|a_n| \leqq \frac{V}{n^2\pi}, \qquad |b_n| \leqq \frac{V}{n^2\pi}.$$

For the case of interpolation, consider once more the sum Σ_0, and its representation by the expression (2). It has already been remarked that $w_{2p} = 0$; let the last term in (2), instead of being dropped, be replaced by

$$[f(x_{2p}) - f(x_{2p+1})]\, w_{2p},$$

with the understanding that $x_{2p+1} = 2\pi$. Then, if the general value of w_q from (4) is substituted in (2), it appears that

$$\Sigma_0 = \frac{1}{1 - e^{[2n\pi i/(2p+1)]}}\left[\sum_{q=0}^{2p}[f(x_q) - f(x_{q+1})] - \sum_{q=0}^{2p}[f(x_q) - f(x_{q+1})]\, e^{nx_{q+1}i}\right].$$

The first sum on the right reduces to $f(x_0) - f(x_{2p+1})$, which is zero, because

* Cf. Picard, loc. cit., pp. 255–256.

of the periodicity of $f(x)$. In the second sum, let the subscript q be replaced by k, and let

$$d_k = f(x_{k+1}) - f(x_k);$$

since $x_{q+1} = x_q + 2\pi/(2p+1)$, the expression for Σ_0 reduces to

$$\Sigma_0 = \frac{e^{[2n\pi i/(2p+1)]}}{1 - e^{[2n\pi i/(2p+1)]}} \sum_{k=0}^{2p} d_k e^{nx_k i}. \tag{5}$$

This formula corresponds to the result of the integration by parts in the case of the Fourier's coefficients.

In the factor before the sign of summation on the right, the absolute value of the numerator is 1. The denominator is the one that was considered in the preceding section; its absolute value is $2 \sin [n\pi/(2p+1)]$, and is greater than $4n/(2p+1)$. Consequently

$$|\Sigma_0| \leqq \frac{2p+1}{4n} \left| \sum_{k=0}^{2p} d_k e^{nx_k i} \right|.$$

By the mean value theorem,

$$d_k = (x_{k+1} - x_k) f'(\xi_k) = \frac{2\pi}{2p+1} f'(\xi_k), \qquad x_k < \xi_k < x_{k+1}.$$

Let a function $\phi(x)$ be defined so as to be equal to $f'(\xi_k)$ for $x = x_k$, $k = 0, 1, \cdots, 2p$, equal to $f'(\xi_0)$ for $x = 2\pi$, linear between any two successive points $2k\pi/(2p+1)$, and periodic with the period 2π. Then $\phi(x)$ is of limited variation in any finite interval, its graph being made up of a finite number of linear segments, and so satisfies the conditions imposed on $f(x)$ at the beginning of the preceding section. Its total variation from $x = 0$ to $x = 2\pi$ is

$$|f'(\xi_1) - f'(\xi_0)| + |f'(\xi_2) - f'(\xi_1)| + \cdots + |f'(\xi_{2p}) - f'(\xi_{2p-1})| + |f'(\xi_0 + 2\pi) - f'(\xi_{2p})|,$$

and this quantity, while not necessarily equal to V, is at any rate not greater than V. Hence the reasoning of the earlier section is applicable, and it appears that

$$\left| \sum_{k=0}^{2p} \phi(x_k) e^{nx_k i} \right| \leqq V \cdot \frac{2p+1}{2n}.$$

From this can be deduced successively the inequalities

$$\left| \sum_{k=0}^{2p} d_k e^{nx_k i} \right| \leqq \frac{V\pi}{n}, \qquad |\Sigma_0| \leqq \frac{(2p+1)V\pi}{4n^2},$$

$$|a_{np}| \leqq \frac{V\pi}{2n^2}, \qquad |b_{np}| \leqq \frac{V\pi}{2n^2}.$$

The conclusion may be stated as follows:*

THEOREM II. *If $f(x)$ is a function of period 2π which has a first derivative $f'(x)$ of limited variation in any finite interval, then*

$$|a_{np}| \leqq \frac{L_2}{n^2}, \qquad |b_{np}| \leqq \frac{L_2}{n^2}, \qquad n > 0,$$

where L_2 is independent of n and p.

If V is the total variation of $f'(x)$ in the interval from 0 to 2π, the constant L_2 can be taken equal to $V\pi/2$.

4. **Proof of convergence.** If $f(x)$ satisfies the hypotheses of the preceding section, the inequalities found there for the coefficients a_n, b_n make it clear that the Fourier's series for $f(x)$ is uniformly convergent. It is not so immediately obvious from first principles that the sum of the series must be $f(x)$. On the other hand, since a_{np} and b_{np} depend on p as well as on n, it is not evident, even as a result of Theorem II, that the interpolating function $\sigma_p(x)$ approaches a definite limit as p becomes infinite; but it is readily seen that if $\sigma_p(x)$ does converge uniformly to a limiting function $g(x)$, then $g(x)$ must be identical with $f(x)$.† More explicitly, if $\sigma_p(x)$ converges to the limit $g(x)$ when p becomes infinite through all integral values, it will converge to the same limit if p is restricted to the values $(3^m - 1)/2$, so that $2p + 1 = 3^m$, $m = 1, 2, 3, \cdots$. But in the latter case, if x is any particular value of the form $2k\pi/(3^\mu)$, where μ is any positive integer and k is any integer whatever, $\sigma_p(x)$ will be equal to $f(x)$ for $m = \mu$ and for all succeeding values of m, because of the properties of $\sigma_p(x)$ as an interpolating function, and therefore the value of $g(x)$ at this point must be the same as that of $f(x)$. That is, $g(x)$ is equal to $f(x)$ at all points $x = 2k\pi/(3^\mu)$; and as both $f(x)$ and $g(x)$ are continuous, and these points are everywhere dense, $g(x)$ must be equal to $f(x)$ identically. It will be shown that $\sigma_p(x)$ converges uniformly to the same limit as $s_p(x)$, and then it will follow, by what has just been said, that both approximating functions converge† to the value $f(x)$.

Let

$$g(x) = \lim_{p=\infty} s_p(x);$$

also, let $\sigma_{np}(x)$ be the sum of the first terms of $\sigma_p(x)$, through those of

* In connection with the value of L_2, cf. § 5 below.

† The truth of this conclusion is well known, for the interpolating function as well as for the Fourier's series, with more general hypotheses as to the function $f(x)$; cf. Faber, loc. cit., and also de la Vallée Poussin, *Sur la convergence des formules d'interpolation entre ordonnées équidistantes*, Bulletins de l'Académie royale de Belgique, Classe des Sciences, 1908, pp. 319–403. The purpose of the present section is to answer the natural question what connection there is between the order of magnitude of the coefficients and the convergence of the interpolating function, and incidentally to call attention to the simplicity of the resulting convergence proof.

order n,

$$\sigma_{np}(x) = \frac{a_{0p}}{2} + \sum_{k=1}^{n} (a_{kp} \cos kx + b_{kp} \sin kx).$$

The convergence of $s_p(x)$, as already noted, follows from the inequalities derived in the preceding section for a_n and b_n. In consequence of this convergence, and of the inequalities for a_{np} and b_{np}, there will exist for every positive ϵ a subscript n such that

$$|g(x) - s_n(x)| < \frac{\epsilon}{3}, \qquad |\sigma_p(x) - \sigma_{np}(x)| < \frac{\epsilon}{3},$$

both inequalities holding for all values of x, and the latter one for all values of $p > n$. If this value of n is held fast, and p is allowed to increase without limit, a_{kp} and b_{kp} will approach a_k and b_k respectively, for all values of $k \leqq n$, $\sigma_{np}(x)$ will uniformly approach $s_n(x)$, and, when p is sufficiently large,

$$|s_n(x) - \sigma_{np}(x)| < \frac{\epsilon}{3}$$

for all values of x. By combination of the last three inequalities, it is seen that

$$|g(x) - \sigma_p(x)| < \epsilon$$

as soon as p is sufficiently large, and this completes the proof of convergence.

5. **Alternative method for a function having a derivative of limited variation.** If the convergence of the Fourier's series to the value $f(x)$, in the case that $f(x)$ has a first derivative of limited variation, be assumed as known at the start, inequalities for the order of magnitude of the coefficients in the interpolating formula can be obtained without the use of partial summation. When $f(x_k)$ in (1) is replaced by its expression in terms of the Fourier's series, it is found that*

$$a_{np} = a_n + \sum_{\lambda=1}^{\infty} (a_{\lambda(2p+1)+n} + a_{\lambda(2p+1)-n}),$$

$$b_{np} = b_n + \sum_{\lambda=1}^{\infty} (b_{\lambda(2p+1)+n} - b_{\lambda(2p+1)-n}).$$

On the other hand, by the inequalities at the beginning of § 3,

$$|a_n| \leqq \frac{K_2}{n^2}, \qquad |b_n| \leqq \frac{K_2}{n^2},$$

where K_2 is a constant. Furthermore, since $n \leqq p$,

$$\lambda(2p+1) + n > (2\lambda+1)n, \qquad \lambda(2p+1) - n > (2\lambda-1)n,$$

* Cf., e.g., de la Vallée Poussin, loc. cit., p. 383.

$$|a_{\lambda(2p+1)+n}| \leqq \frac{K_2}{(2\lambda+1)^2 n^2}, \qquad |b_{\lambda(2p+1)+n}| \leqq \frac{K_2}{(2\lambda+1)^2 n^2},$$

$$|a_{\lambda(2p+1)-n}| \leqq \frac{K_2}{(2\lambda-1)^2 n^2}, \qquad |b_{\lambda(2p+1)-n}| \leqq \frac{K_2}{(2\lambda-1)^2 n^2},$$

so that

$$|a_{np}| \leqq \frac{2K_2}{n^2}\left(1+\frac{1}{3^2}+\frac{1}{5^2}+\cdots\right) = \frac{L_2'}{n^2}, \qquad |b_{np}| \leqq \frac{L_2'}{n^2},$$

L_2' being independent of n.

Since*

$$1+\frac{1}{3^2}+\frac{1}{5^2}+\cdots = \frac{\pi^2}{8},$$

and since, by the preceding section, K_2 may be taken equal to V/π, the value $V\pi/4$ is found for L_2', which is an improvement over the value of L_2 given in the statement of Theorem II.

If $f(x)$ possesses a derivative of order higher than the first having limited variation, the first derivative is certainly of limited variation, and it follows from § 4, even if it is not assumed otherwise, that the Fourier's series converges to the value $f(x)$. The method just outlined can be used, with the stronger inequalities for a_n and b_n which are obtained by integration by parts in this case, to yield corresponding inequalities for a_{np} and b_{np}.

6. **The method of partial summation for a function having a second derivative of limited variation.** Although the existence of higher derivatives can be taken into account in the manner indicated at the end of the preceding section, it is not without interest to observe that the method of §§ 2, 3 can be made to apply to this case also. The extension is not entirely trivial, but it will perhaps be sufficient to carry it one stage further. Let $f''(x)$, then, be everywhere defined, and of limited variation in any finite interval. It is to be shown that

$$|a_{np}| \leqq \frac{L_3}{n^3}, \qquad |b_{np}| \leqq \frac{L_3}{n^3},$$

where L_3 is independent of n and p. The more exact determination of the constant is of secondary interest.

Let it be understood now that

$$x_k = \frac{2k\pi}{2p+1},$$

for all integral values of k, whether contained in the interval from 0 to $2p+1$ or not; let

$$d_k' = f(x_{k+1}) - f(x_k) = d_k,$$

* Cf., e.g., Bromwich, *Theory of infinite series*, p. 187.

an accent being added to the symbol d_k previously used, to indicate that these are first differences, as distinguished from the second differences now to be introduced; and let

$$d_k'' = d_{k+1}' - d_k'.$$

The coefficients a_{np} and b_{np}, apart from the factor $2/(2p+1)$, are the components of the complex quantity Σ_0 already considered. If

$$\Sigma_1 = \sum_{k=0}^{2p} d_k' e^{nx_k i},$$

it follows from (5) that

$$\Sigma_0 = \frac{e^{[2n\pi i/(2p+1)]}}{1 - e^{[2n\pi i/(2p+1)]}} \Sigma_1.$$

The sum Σ_1 can be rearranged by partial summation, after the same manner as Σ_0. It is found that

$$\begin{aligned}\Sigma_1 &= d_0' w_0 + d_1'(w_1 - w_0) + \cdots + d_{2p}'(w_{2p} - w_{2p-1}) \\ &= (d_0' - d_1') w_0 + (d_1' - d_2') w_1 + \cdots + (d_{2p-1}' - d_{2p}') w_{2p-1} + d_{2p}' w_{2p};\end{aligned}$$

since $w_{2p} = 0$, the last term may be replaced by $(d_{2p}' - d_{2p+1}') w_{2p}$. In consequence of (4), furthermore,

$$\Sigma_1 = \frac{1}{1 - e^{[2n\pi i/(2p+1)]}} \left[\sum_{q=0}^{2p} (d_q' - d_{q+1}') - \sum_{q=0}^{2p} (d_q' - d_{q+1}') e^{nx_{q+1}i} \right].$$

The first sum on the right reduces to $d_0' - d_{2p+1}'$, and this is equal to zero, because d_{2p+1}' is the same as d_0'. Hence

$$\Sigma_1 = \frac{e^{[2n\pi i/(2p+1)]}}{1 - e^{[2n\pi i/(2p+1)]}} \sum_{k=0}^{2p} d_k'' e^{nx_k i},$$

and

$$\Sigma_0 = \left(\frac{e^{[2n\pi i/(2p+1)]}}{1 - e^{[2n\pi i/(2p+1)]}} \right)^2 \sum_{k=0}^{2p} d_k'' e^{nx_k i}.$$

It is clear that a corresponding relation can be obtained for differences of any order. By reason of the inequality previously used for the factor in parentheses,

$$|\Sigma_0| \leqq \left(\frac{2p+1}{4n} \right)^2 \left| \sum_{k=0}^{2p} d_k'' e^{nx_k i} \right|.$$

It remains to discuss the order of magnitude of the sum on the right.

By one more partial summation, which it is not necessary to give in detail, it is seen that

$$\left| \sum_{k=0}^{2p} d_k'' e^{nx_k i} \right| \leqq (|d_1'' - d_0''| + |d_2'' - d_1''| + \cdots + |d_{2p}'' - d_{2p-1}''|) H,$$

where H, as in § 2, is an upper bound for $|w_q|$, and may be taken equal to $(2p+1)/(2n)$. So, if X stands for the quantity in parentheses,

$$|\Sigma_0| \leqq \frac{(2p+1)^3}{32n^3} X, \tag{6}$$

and the problem is further reduced to that of treating the expression X.

If a function $f(x)$, possessing a second derivative throughout an interval $a \leqq x \leqq c$, vanishes for $x = a$ and for $x = c$, and also at an intermediate point $x = b$, the first derivative $f'(x)$ must vanish at some point between a and b, and again between b and c, and so $f''(x)$ must vanish somewhere between a and c. If $f(x)$ is not required to vanish at the points a, b, and c, but $A + Bx + Cx^2$ is the polynomial of the second degree agreeing in value with $f(x)$ at these three points, the second derivative of

$$f(x) - (A + Bx + Cx^2)$$

must vanish somewhere between a and c, or, in other words, there is a point at which $f''(x) = 2C$. Now let it be assumed further that b is midway between a and c, so that $b = a + h$, $c = a + 2h$, say. Then the second difference of $f(x)$ for the three points,

$$f(x + 2h) - 2f(x + h) + f(x),$$

is the same as the corresponding second difference formed for the function $A + Bx + Cx^2$, and the latter second difference is equal to $2Ch^2$. That is, since $2C = f''(\xi)$, where ξ has some value between a and c, the second difference of $f(x)$ is equal to* $h^2 f''(\xi)$.

In the problem under discussion, where $h = 2\pi/(2p+1)$,

$$d_k'' = \frac{4\pi^2}{(2p+1)^2} f''(\xi_k),$$

where ξ_k is some point between x_k and x_{k+2}. Consequently

$$X = \frac{4\pi^2}{(2p+1)^2}[|f''(\xi_1) - f''(\xi_0)| + |f''(\xi_2) - f''(\xi_1)| + \cdots + |f''(\xi_{2p}) - f''(\xi_{2p-1})|].$$

If it could be said that $\xi_0 < \xi_1 < \xi_2 < \cdots$, the quantity in brackets could be brought into immediate connection with the total variation of $f''(x)$; but it is not clear that the ξ's do occur exactly in the order of their subscripts. However, since $x_0 < \xi_0 < x_2$, $x_1 < \xi_1 < x_3$, etc., there is no doubt that the pair of points (ξ_0, ξ_1) precedes the pair (ξ_3, ξ_4), etc., and as the order of the terms in any one difference $|f''(\xi_{k+1}) - f''(\xi_k)|$ is immaterial, it can be

* Cf. Markoff, *Differenzenrechnung* (Leipzig, Teubner, 1896), p. 21.

asserted that for present purposes the points $\xi_0, \xi_1, \xi_3, \xi_4, \xi_6, \xi_7, \cdots$, are essentially in order of algebraic magnitude. Moreover, while there may be a question whether the points $\xi_0, \cdots, \xi_{2p}$ are all included in an interval of length 2π, they are at any rate all included in an interval of length 4π. So, if V is the total variation of $f''(x)$ in an interval of length 2π, it is certain that $|f''(\xi_1) - f''(\xi_0)| + |f''(\xi_4) - f''(\xi_3)| + |f''(\xi_7) - f''(\xi_6)| + \cdots \leqq 2V$. Similarly,

$$|f''(\xi_2) - f''(\xi_1)| + |f''(\xi_5) - f''(\xi_4)| + \cdots \leqq 2V,$$

$$|f''(\xi_3) - f''(\xi_2)| + |f''(\xi_6) - f''(\xi_5)| + \cdots \leqq 2V.$$

It would be possible to do something toward replacing these by a stricter set of inequalities, but that point will not be pressed. By combination of the three as they stand,

$$|f''(\xi_1) - f''(\xi_0)| + |f''(\xi_2) - f''(\xi_1)| + \cdots \leqq 6V,$$

$$X \leqq \frac{24\pi^2 V}{(2p+1)^2}.$$

With (6), this gives

$$|\Sigma_0| \leqq \frac{3\pi^2 V}{4} \cdot \frac{2p+1}{n^3}.$$

Finally, because of the factor $2/(2p+1)$ in a_{np} and b_{np},

$$|a_{np}| \leqq \frac{3\pi^2 V}{2n^3}, \qquad |b_{np}| \leqq \frac{3\pi^2 V}{2n^3},$$

the factor $3\pi^2 V/2$ being independent of n and p.

There is no further essential difficulty in carrying the discussion on to derivatives of higher order.

THE UNIVERSITY OF MINNESOTA,
MINNEAPOLIS, MINN.

CONCERNING SIMPLE CONTINUOUS CURVES*

BY

ROBERT L. MOORE

1. Introduction

Various definitions of simple continuous arcs and closed curves have been given.† The definitions of arcs usually contain the requirement that the point-set in question should be bounded. In attempting to prove that every interval t of an open curve as defined in a recent paper‡ is a simple continuous arc, while I found it easy to prove that t satisfies all the other requirements of Janiszewski's definition (modified as indicated below) it was only by a rather lengthy and complicated argument that I succeeded in proving that it satisfies the requirement of boundedness. In Lennes' definition the requirement of boundedness is superfluous.§ However I found it difficult to prove that t satisfies a certain one of the other requirements of this definition, namely that the point-set in question should contain no proper connected subset that contains both A and B. In the present paper I will give a definition‖ of a simple continuous arc which stipulates neither that the set M should be bounded nor that it should contain no proper connected subset containing both A and B. I will show that, in a euclidean space of two dimensions, every point-set that satisfies this definition is an arc in the sense of Jordan. It is easy to prove¶ that every interval of an open curve satisfies this definition.

* Presented to the Society, October 26, 1918.

† Cf., for example, the following: S. Janiszewski, *Sur les continus irréductibles entre deux points*, Journal de l'Ecole Polytechnique, 2e série, vol. 16 (1911–12), pp. 79–170. N. J. Lennes, *Curves in non-metrical analysis situs with an application in the calculus of variations*, American Journal of Mathematics, vol. 33 (1911), p. 308 and Bulletin of the American Mathematical Society, vol. 12 (1906), p. 284. W. Sierpinski, *L'arc simple comme un ensemble de points dans l'espace á m dimensions*, Annali di Matematica, Serie III, vol. 26 (1916), pp. 131–150. J. R. Kline, *Concerning the relation between approachability and connectivity in kleinem.* R. L. Moore, *A characterization of Jordan regions by properties having no reference to their boundaries*, Proceedings of the National Academy of Sciences, vol. 4 (1918), pp. 364–370.

‡ R. L. Moore, *On the foundations of plane analysis situs*, these Transactions, vol. 17 (1916), p. 159.

§ Cf. G. H. Hallett, Jr., *Concerning the definition of a simple continuous arc*, Bulletin of the American Mathematical Society, vol. 25 (1919), pp. 325–326.

‖ Definition 1 below.

¶ See proof of Theorem 3 below.

In connection with certain problems where the boundedness of the point-set in question is not presupposed, but where relatively more information is at hand concerning *connectedness*, it seems likely that Definition 1 may be more useful than that of Janiszewski. On the other hand Janiszewski's definition (or a modification of it*) may be of more use in certain cases where one is concerned with sets that are known in advance to be bounded (or can easily be proved to be bounded) but concerning which less is known in advance with regard to connectedness.†

In the latter part of § 2, a very simple characterization of a simple closed curve is given. It is defined merely as a closed connected and bounded point-set which is disconnected by the omission of any two of its points.‡

In § 3 the problem of defining simple continuous arcs, closed and open curves and rays is approached from a different point of view. With the use of the notion of the boundary of a point-set M with respect to a point-set that contains M, conditions are given which a point-set must satisfy in order that it should be a *simple continuous curve* (that is to say one of the four types of curves mentioned above). This classification of the general notion *simple continuous curve* having been given, it is easy to so particularize it as to obtain a characterization of any given one of the four special types of simple continuous curves. A definition of an open curve from this point of view can be obtained from that of a closed curve by the mere substitution of the word "bounded" in place of the word "proper."

2

DEFINITION 1. If A and B are two distinct points, a simple continuous arc from A to B is a closed, connected set of points M containing A and B such that (1) $M - A$ and $M - B$ are connected, (2) if P is any point of M

* Cf. Definition 2 below.

† In this connection see the article by Sierpinski referred to above. His definition does not require explicitly even that M itself should be connected. It requires boundedness however and also certain positive information concerning the relation of the endpoints A and B to the set M. In my Definition 2 the phrase "except the points A and B" is to be interpreted not as meaning that the points A and B do not fulfill the requirements indicated but merely as leaving the question open whether they do or do not fulfill these requirements. Sierpinski definitely stipulates that M is not the sum of two closed point-sets having only A, or only B, in common and each consisting of more than one point. If this stipulation were omitted his definition would not completely characterize a simple continuous arc and would indeed apply to some sets that are not connected, e.g., to a set composed of three distinct points.

‡ Lennes defines a simple closed curve as "the set of points consisting of two continuous arcs, each connecting a pair of distinct points A and B and having no other point in common." This definition presupposes a previous definition of a simple continuous arc. Janiszewski (loc. cit., p. 137) defines "*Une ligne simple fermée* Γ" as "un continu qui peut être décomposé en deux continus c_1 et c_2 n'ayant en commun que deux points M et N arbitrairement donnés sur Γ."

distinct from A and from B then $M - P$ is the sum of two mutually exclusive connected point-sets neither of which contains a limit point of the other one.

If for a point O of a connected point-set M the set $M - O$ is the sum of two mutually exclusive point-sets M_1 and M_2 neither of which contains a limit point of the other one then M_1 and M_2 will be called sects* (of M) from O, and M will be said to be disconnected by the omission of O and will be said to be separated by the omission of O into the two sets M_1 and M_2. If P is a point of M distinct from O then in case there is only one sect of M from O that contains P that sect will be called the sect OP. If at the same time there is only one sect from P that contains O the set of all those points of M that are common to the sects OP and PO will be called the *segment* OP (of M) while the set of points consisting of all the points of the segment OP together with its *endpoints* (O and P) will be called the *interval* OP (of M).

THEOREM 1. *In a Euclidean space of two dimensions every set of points M that satisfies the requirements of Definition* 1 *is an arc in the sense of Jordan.*

Proof. Suppose M is a set of points satisfying all the requirements of Definition 1. If O is any point of M distinct from A and from B it is easy to see that there are only two sects of M from O. One of these sects contains A and the other one contains B. For suppose that one of them contains both A and B. Let OC denote the other one. Let K denote the set of all points $[X]$ such that one sect from X contains neither A nor B but has at least one point in common with OC. For every point P of K let K_P denote that sect from P which contains neither A nor B and let $\overline{K}_P$ denote the other sect from P.

The points of OC can† be arranged in a well-ordered sequence β. Let P_1 be the first point in the sequence β. Let P_2 be the first point of β which lies in K_{P_1}. Let P_3 be the first point of β which is common to K_{P_1} and K_{P_2}. Let P_4 be the first point of β which is common to K_{P_1}, K_{P_2}, and K_{P_3}. This process may be continued. It follows that the sequence β contains a well-ordered subsequence α such that if T is a subset of the elements of α there is an element of α which follows, in α, all the elements of T if, and only if, there exists a point which belongs to K_P for every point P of T and, if there does exist such a point, then the first element of α that follows all the elements of T is the first element of β which belongs to K_P for every point P of T. Suppose that X and Y are distinct points of α and that Y precedes X in α. Then X is in K_Y and therefore is not in $\overline{K}_Y + Y$ and hence $\overline{K}_Y + Y$ must

* The term "sect" is used by Halsted with a somewhat different meaning. Cf. G. B. Halsted, *Rational Geometry*, Wiley and Sons, New York, 1904.

† The Zermelo Postulate is here assumed. Cf. E. Zermelo, *Beweiss, dass jede Menge wohlgeordnet sein kann*, Mathematische Annalen, vol. 59 (1904), pp. 514–516. Concerning this postulate cf. Philip E. B. Jourdain, Comptes Rendus, vol. 166 (1918), pp. 520–523 and 984–986.

lie wholly in K_X or wholly in $\bar{K}_X$ (otherwise $\bar{K}_Y + Y$ would not be connected and consequently $K_Y + Y + \bar{K}_Y$ would not be connected). But $\bar{K}_Y$ contains A and K_X does not. Hence $\bar{K}_Y + Y$ is a subset of $\bar{K}_X$. Therefore K_X contains no point of $\bar{K}_Y + Y$. It follows that K_X is a subset of K_Y. But of any two points in α one of them precedes the other one in α. Hence if X and Y are two points of α either K_X contains K_Y or K_Y contains K_X.

If X and Y are distinct points of K and K_X contains K_Y then K_Y does not contain K_X. For since $X + K_X$ is closed, $X \neq Y$ and Y is a limit point of K_Y, therefore Y is in K_X. But Y is not in K_Y. Hence K_Y does not contain K_X.

If, for two distinct points X and Y of the set K, the sect K_X contains the sect K_Y then X will be said to precede Y in K. In view of the results established above it is clear that if X and Y are two distinct points in K then (1) if X precedes Y, Y does not precede X, (2) if X precedes Y and Y precedes Z then X precedes Z, (3) if X and Y are both elements of α then either X precedes Y or Y precedes X.

The ray OC is unbounded. For suppose it is bounded. Then the family of all sets $K_X + X$ for all points X of α is a family of closed, bounded point-sets such that of every two of them one contains the other one. It follows by a theorem established in a recent paper* that there exists at least one point W which belongs to $K_X + X$ for every point X of α. Let W_1 denote the first such point W in the sequence β. Then W_1 is an element of α that follows all the elements of α. Thus the supposition that OC is bounded leads to a contradiction.

Suppose that X and Y are two distinct points of K that do not belong to $O + OC$. Since the connected set $O + OC$ contains a point of K_X but does not contain X therefore $O + OC$ is a subset of K_X. Similarly $O + OC$ is a subset of K_Y. Thus K_X and K_Y have at least one point in common. Suppose now that K_X is not a subset of K_Y. Then since $X + K_X$ is connected and contains at least one point of K_Y, K_X must contain Y. Hence $X + \bar{K}_X$ does not contain Y. But $\bar{K}_X + X$ is connected and contains the point A in common with $\bar{K}_Y$. Hence $\bar{K}_X + X$ is a subset of $\bar{K}_Y$ and therefore $K_Y + Y$ is a subset of K_X. Thus it is proved that if X and Y are two distinct points of $K - OC - O$ then either X precedes Y or Y precedes X.

Let H denote the set of points composed of K together with the set of all points $[Y]$ such that Y belongs to K_X for some point X of K.

Since M is connected either H contains a limit point of $M - H$ or $M - H$ contains a limit point of H. Suppose that $M - H$ contains a point Z which

* R. L. Moore, *On the most general class L of Fréchet in which the Heine-Borel-Lebesgue theorem holds true*, Proceedings of the National Academy of Sciences, vol. 5 (1919), pp. 206–210. Cf. also S. Janiszewski, loc. cit.

is a limit point of H. Then Z is the sequential limit point of some infinite sequence of distinct points $Z_1, Z_2, Z_3, \cdots$ belonging to H. It follows that there exists an infinite sequence of distinct points $X_1, X_2, X_3, \cdots$ belonging to $K - O - OC$ and an infinite sequence of distinct positive integers $n_1, n_2, n_3, \cdots$ such that, for every m, (1) Z_{n_m} belongs to K_{X_m} and (2) X_{m+1} precedes X_m in K. Let K^* denote the point-set $X_1 + X_2 + X_3 + \cdots$. It is clear that (1) if Y is a point of H there exists a positive integer m such that K_{X_m} contains Y, (2) $M - H$ contains every limit point of K^*. Every limit point of H which lies in $M - H$ is a limit point of K^*. Suppose $M - H$ contains two points D and E which are limit points of H. Then it can be shown that there exist in K^* two sequences of points $D_1, D_2, D_3, \cdots$ and $E_1, E_2, E_3, \cdots$ such that (1) D and E are sequential limit points of $D_1, D_2, D_3, \cdots$ and of $E_1, E_2, E_3, \cdots$ respectively, (2) for every i, E_i is preceded by D_i and D_i is preceded by E_{i+1} in K. There exist five regions§ R_1, R_2, R_3, R_4, and R_5 such that (1) R_3 contains D and R_4 contains E, (2) R_3' is a subset of R_2, R_1' is a subset of R_1 and R_1' is a subset of R_5, (3) R_5' and R_2' have no point in common. There exists a positive integer n such that for every positive integer m the point D_{n+m} is in R_3 and the point E_{n+m} is in R_4. It is clear that the intervals $E_{n+1}D_{n+1}$, $E_{n+2}D_{n+2}$, $E_{n+3}D_{n+3}$, $\cdots$ of M are all closed connected subsets of K and that no two of them have a point in common. For every m the interval $E_{n+m}D_{n+m}$ contains a closed connected subset t_m that contains at least one point on the boundary of R_1 and at least one point on the boundary of R_4 but contains no point without R_1 or within R_4. No two point-sets belonging to the infinite sequence $t_1, t_2, t_3, \cdots$ have a point in common.† It follows that there exists (1) an infinite sequence of positive integers $n_1, n_2, n_3, \cdots$ such that, for every j, $n_{j+1} > n_j$ and (2) a closed connected set of points t and a sequence of closed connected point-sets $k_{n_1}, k_{n_2}, k_{n_3}, \cdots$ such that, for every j, k_{n_j} is a subset of t_{n_j} and such that (a) each of the point-

§ In my paper *On the foundations of plane analysis situs*, loc. cit. (this paper will be referred to as F.A.), the notion *point* is undefined and *region* is also undefined except in so far as it is understood that every region is some sort of collection of points. A considerable part of the present argument holds good not only for a euclidean space of two dimensions but for any space satisfying Axioms 1′ and 4 of F.A. The whole of this argument holds good for every space satisfying the set of Axioms Σ_1 of F.A. Every such space is, however, in one to one continuous correspondence with an ordinary euclidean space of two dimensions; cf. my paper *Concerning a set of postulates for plane analysis situs*, these Transactions, vol. 20 (1919), pp. 169–178. If one desires to read the present paper without special reference to any particular system of axioms he may think of the word *region* as applying to any bounded connected domain in a euclidean space of two dimensions.

† Up to this point the present proof holds good for every space satisfying Axioms 1′ and 4 of F.A. It therefore holds good for all euclidean spaces (of however many dimensions) as well as for many other spaces including certain spaces that are neither metrical, descriptive, nor separable (cf. F.A., p. 131). If the statement in the next sentence (which can be easily established for euclidean space of two dimensions) can be proved to hold true in euclidean space of any number of dimensions then the present proof holds good for any such space.

sets $t, k_{n_1}, k_{n_2}, k_{n_3}, \cdots$ is a subset of $R_1' - R_4$ but contains at least one point on the boundary of R_1' and at least one point on the boundary of R_4, (*b*) if $P_{n_1}, P_{n_2}, \cdots$ is a sequence of points such that, for every j, P_{n_j} belongs to k_{n_j} then t contains every limit point of the point-set $P_{n_1} + P_{n_2} + P_{n_3} + \cdots$, (*c*) if $n_1, \bar{n}_2, \bar{n}_3, \cdots$ is an infinite sequence of distinct integers belonging to the set $n_1, n_2, n_3, \cdots$ and P is a point of t there exists an infinite sequence of points $P_{\bar{n}_1}, P_{\bar{n}_2}, P_{\bar{n}_3}, \cdots$ such that, for every j, $P_{\bar{n}_j}$ belongs to $k_{\bar{n}_j}$ and such that P is the sequential limit point of the sequence $P_{\bar{n}_1}, P_{\bar{n}_2}, P_{\bar{n}_3}, \cdots$. If T is a point of t, T is the sequential limit point of some sequence of points $T_{n_1}, T_{n_2}, T_{n_3}, \cdots$ such that, for every j, T_{n_j} belongs to k_{n_j}. Of the two sects of M from T one must contain an infinite subsequence $T_{n_{j_1}}, T_{n_{j_2}}, T_{n_{j_3}}, \cdots$ of the sequence $T_{n_1}, T_{n_2}, T_{n_3}, \cdots$. For every positive integer m the connected point-set $k_{n_{j_m}}$ contains $T_{n_{j_m}}$ but does not contain T. It follows that that sect from T which contains the points $T_{n_{j_1}}, T_{n_{j_2}}, T_{n_{j_3}}, \cdots$ contains also the point-sets $k_{n_{j_1}}, k_{n_{j_2}}, k_{n_{j_3}}, \cdots$. But every point of t is a limit point of $k_{n_{j_1}} + k_{n_{j_2}} + \cdots$. Hence that sect contains every point of $t - T$. For every point T of t let $\bar{R}_T$ denote that sect from T which contains $t - T$ and let R_T denote the other sect from T. If T_1 and T_2 are two distinct points of t, $\bar{R}_{T_1}$ contains the point T_2 of the connected point-set $R_{T_2} + T_2$ and T_1 does not belong to $R_{T_2} + T_2$. It follows that R_{T_2} is a subset of $\bar{R}_{T_1}$. Hence R_{T_1} and R_{T_2} have no point in common. Hence there do not exist more than two points X belonging to t such that R_X contains A or B. Let t_0 denote t, $t - X_1$, or $t - (X_1 + X_2)$ according as there are no such points X or there is only one such point X_1 or there are two such points X_1 and X_2. For each point T of t_0 the sect R_T is unbounded (cf. above) and contains a point (T) in R_1'. Hence it contains at least one point P_T on the boundary of R_5. Consider the set L of all P_T's for all points T of t_0. There is a one to one correspondence between the point-sets L and t_0. It follows that L is uncountable. But no point of any sect R_T is a limit point of $M - R_T$. Hence no point of L is a limit point of L. But every uncountable set of points α contains at least one point which is a limit point of α. Thus the supposition that $M - H$ contains more than one limit point of H has led to a contradiction.

It is clear that H cannot contain more than one limit point of $M - H$. For no point of OC is a limit point of $M - OC$ and if X and Y are two points of H not belonging to OC there exist two points $\bar{X}$ and $\bar{Y}$ belonging to K such that $\bar{X} + K_{\bar{X}}$ contains X and $K_{\bar{Y}} + \bar{Y}$ contains Y. But either $K_{\bar{X}}$ contains $K_{\bar{Y}} + \bar{Y}$ or $K_{\bar{Y}}$ contains $\bar{K}_{\bar{X}} + \bar{X}$. In the first case Y is not a limit point of $\bar{K}_{\bar{X}}$ and therefore is not a limit point of $M - H$ which is a subset of $\bar{K}_X$. In the second case X is not a limit point of $M - H$.

It follows that there exists one and only one point O which belongs to one of the sets H and $M - H$ and is a limit point of the other one.

Suppose that O belongs to $M - H$. Neither of the sets H and $M - (H + O)$ contains a limit point of the other one. But K is connected and each sect from O is connected. It follows that K is one of the sects of M from O. Thus the supposition that O belongs to $M - H$ leads to a contradiction. Hence O belongs to H. Thus for every point P of M such that one sect from P contains neither A nor B there exists a point O_P such that (1) the sect $O_P P$ contains neither A nor B, (2) the sect $O_P P$ is not a subset of any other sect that contains neither A nor B. Let $\overline{H}$ denote the set of all points P of the segment AB such that one sect from P contains neither A nor B. Let N_1 denote the set of all points O_P for all points P of $\overline{H}$. For each point X of $\overline{H}$ let M_X denote that sect from X which contains neither A nor B. Let N_2 denote the set of all those points of M that belong neither to N_1 nor to any M_X for any point X of N_1. Let N denote the set $N_1 + N_2$. Every point of N_1 will be called an improper point and every point of N_2 will be called a proper point. Either N_2 contains a limit point of $\overline{H}$ or $\overline{H}$ contains a limit point of N_2. Suppose first that N_2 contains a point O which is a limit point of $\overline{H}$. Then there exists a sequence of distinct points $X_1, X_2, X_3, \cdots$ belonging to N_1 and a sequence of points $P_1, P_2, P_3, \cdots$ such that O is the sequential limit point of the sequence $P_1, P_2, P_3, \cdots$ and such that, for every n, P_n either coincides with X_n or belongs to M_{X_n}. There exist about O two regions $\overline{R}_1$ and $\overline{R}_2$ such that $\overline{R}_2'$ is a subset of $\overline{R}_1$. There exists a positive integer n such that the points $P_n, P_{n+1}, P_{n+2}, \cdots$ are all in $\overline{R}_2$. But each of the point-sets $P_n + M_{P_n}$, $P_{n+1} + M_{P_{n+1}}$, $P_{n+2} + M_{P_{n+2}}$, $\cdots$ is closed, connected, and unbounded. It follows that for each m, $M_{P_{n+m}}$ contains a closed, connected subset $\bar{t}_m$ which contains at least one point on the boundary of $\overline{R}_1$ and at least one point on the boundary of $\overline{R}_2$ and every point of which is either on the boundary of $\overline{R}_1$ or of $\overline{R}_2$ or in the domain $\overline{R}_1 - \overline{R}_2'$. No two of the point-sets $\bar{t}_1, \bar{t}_2, \bar{t}_3, \cdots$ have a point in common. That this leads to a contradiction follows by an argument analogous to (or identical with) that used in a similar connection above.

Suppose secondly that $\overline{H}$ contains a point O which is a limit point of N_2. Clearly O must belong to N_1. If X is a point of N_2 distinct from A and from B and Y is a point belonging either to N_1 or to N_2, X will be said to precede Y or to follow Y according as Y belongs to the sect XB or to the sect XA. It may be easily proved that there exists in N_2 a sequence of distinct points $X_1, X_2, X_3, \cdots$, all distinct from A and from B, such that O is a sequential limit point of this sequence and such that either (1) for every n, X_n precedes O and X_{n+1} or (2) for every n, X_n follows O and X_{n+1}. Suppose that, for every n, X_n precedes O and X_{n+1}. Each of the intervals $X_1X_2, X_2X_3, \cdots$ is closed and connected, no two of them contain in common any point other than a common endpoint and no one of them contains the point O. That O is

the only limit point of $X_1X_2 + X_2X_3 + \cdots$ that does not belong to $X_1X_2 + X_2X_3 + \cdots$ may be proved by an argument similar to that used above in the proof that $M - H$ does not contain more than one limit point of H.

Now sect

$$OA = M - M_O - O = (\overline{AX_1}^* + \overline{X_1X_2} + \overline{X_2X_3} + \cdots) + (B + V)$$

where V is the point-set $M - (B + M_O + O + A\overline{X_1} + \overline{X_1X_2} + \overline{X_2X_3} + X_3\overline{X_4} + \cdots)$. Neither of the two sets $(A\overline{X_1} + X_1X_2 + \overline{X_2X_3} + \cdots)$ and $(B + V)$ contains a limit point of the other one. Hence the sect OA is not connected. This involves a contradiction. Hence it is not true that, for every n, X_n precedes O and X_{n+1}. In an entirely similar way it may be proved that X_n cannot follow O and X_{n+1} for every n. Thus the supposition that $\overline{H}$ contains a limit point of N_2 has led to a contradiction.

It follows that the set $\overline{H}$ does not exist. Hence if P is any point of $M - (A + B)$, $M - P$ is the sum of two sects of which one contains A and the other contains B. It follows that the two sects PA and PB have no point in common and neither of them contains a limit point of the other one. Suppose now that $\overline{M}$ is a proper subset of M that contains both A and B. Then M contains a point P that does not belong to $\overline{M}$. Let $\overline{M}_1$ denote the set of all points common to $\overline{M}$ and the sect PA and let $\overline{M}_2$ denote the set of all points common to $\overline{M}$ and the sect PB. Neither of the sets $\overline{M}_1$ and $\overline{M}_2$ contains a limit point of the other one. But $\overline{M} = \overline{M}_1 + \overline{M}_2$. Hence $\overline{M}$ is not connected. Thus M contains no proper connected subset that contains both A and B. It follows that M satisfies all the requirements of Lennes' definition of an arc with the exception of the requirement that it should be bounded. That it also satisfies the latter requirement follows from the theorem of Hallett referred to above. The truth of Theorem 1 is therefore established.

DEFINITION 2.† If A and B are two distinct points a *simple continuous arc* from A to B is a closed, connected, and bounded point-set containing A and B which is disconnected by the omission of any one of its points which is distinct from A and from B.

THEOREM 2. *In a space satisfying Axioms*‡ 1′ *and* 4 *of F.A. every point-*

* The notation $\overline{AB}$ will be used to denote the interval AB of M.

† This definition is closely related to those of Janiszewski and Sierpinski, loc. cit. Janiszewski's definition contains an unnecessary requirement concerning connectedness and also the requirement that the point-set M should be an "irreducible continu from A to B" i.e., that it should contain no proper closed and connected subset containing both A and B. He indicates that this latter condition is redundant and makes reference in this connection to a proof of another theorem. I have not succeeded however in seeing that the argument given there proves the redundancy in question. As I have already observed, Sierpinski's definition contains a certain stipulation concerning A and B.

‡ Loc. cit., pp. 163 and 132.

set M that satisfies the requirements of Definition 2 satisfies also those of Lennes' definition.

Proof. Let M be a set of points satisfying the requirements of Definition 2. Certain portions of the proof of Theorem 1 clearly apply here. In particular the same argument that was used there applies here to show that if P is any point of M distinct from A and from B and $M - P$ is the sum of two subsets neither of which contains a limit point of the other one then if one of these subsets contains neither A nor B it must be unbounded. But here M itself is bounded. Therefore, for every point P of $M - (A + B)$, $M - P$ is* the sum of two mutually separated† point-sets M_A and M_B such that M_A contains A and M_B contains B. It follows‡ that M contains no proper connected subset containing both A and B. The truth of Theorem 2 is therefore established.

DEFINITION 3. An *open curve* is a closed and connected point-set which is separated into two connected subsets by the omission of any one of its points.

If P is a point of an open curve M the point-set obtained by adding P to either of the two sects into which M is separated by the omission of P is called a *ray*. The two rays of M so determined by a point P are said to start from P. If A is a point of M distinct from P that ray of M which starts from P and contains A will be called the ray PA.

THEOREM 3. *In euclidean space of two dimensions, if A and B are two points of the open curve M the interval AB of M is a simple continuous arc from A to B.*§

Proof. It is clear that the interval $\overline{AB}$ is closed. It is connected. For suppose it is not. Then it is the sum of two mutually separated point-sets. Let M_A denote that one of these sets which contains A and let $\overline{M}_A$ denote the other one. Let AC denote that ray from A which does not contain B and let BD denote that ray from B which does not contain A. If $\overline{M}_A$ contains B then neither of the complementary sets $AC + M_A$ and $\overline{M}_A + BD$ contains a limit point of the other one. If $\overline{M}_A$ does not contain B then neither of the complementary sets $AC + BD + M_A$ and $\overline{M}_A$ contains a limit point of the other one. Thus in either case M is not connected. Thus the supposition that $\overline{AB}$ is not connected has led to a contradiction.

* Cf. an argument given by Sierpinski, loc. cit. pp. 137–140. His argument assumes separability while the proof given here holds good in every space satisfying Axioms 1′ and 4. Such spaces are of course not necessarily separable.

† Two point-sets are said to be *mutually separated* if neither contains a point or a limit point of the other one.

‡ See latter part of proof of Theorem 1.

§ The definition of an open curve given on page 159 of F.A. is (aside from phraseology) equivalent to Definition 3. Theorem 49 is however not fully proved in F.A. In particular it is there assumed without proof that t contains no proper connected subset that contains both A and B.

Suppose that O is any point of the interval $\overline{AB}$ distinct from A and from B. It will be shown that $\overline{AB} - O$ is the sum of two connected point-sets neither of which contains a limit point of the other one. Suppose that one of the sects of M from O contains both A and B. Let OX denote the other sect of M from O. Let N denote the set of all those points which are common to the sect OA and the interval $\overline{AB}$. The point-set N contains A and B. Suppose it is not connected. Then it is the sum of two sets neither of which contains a limit point of the other one. Let N_A denote that one of these sets which contains A and let N_B denote the other one. If N_A contains B then the sect OA is the sum of the two mutually separated sets $AC + BD + N_A$ and N_B. If N_A does not contain B then the sect OA is the sum of the two mutually separated sets $AC + N_A$ and $BD + N_B$. In either case the sect OA is not connected. Thus the supposition that one sect of M from O contains both A and B leads to a contradiction. It follows that the sects OA and OB have no point in common. Hence neither of the point-sets $\overline{OA} - O$ and $\overline{OB} - O$ contains a point or a limit point of the other one. It is easy to see that these point-sets are connected and that their sum is $\overline{AB} - O$. Thus if O is any point of the interval AB distinct from A and from B then $\overline{AB} - O$ is the sum of two connected point-sets neither of which contains a point or a limit point of the other one. Thus the interval $\overline{AB}$ satisfies all the requirements for an arc as given in Definition 1.

For further results concerning open curves see F.A., loc. cit., and a paper by J. R. Kline.*

DEFINITION 4. A simple closed curve is a closed, connected, and bounded point-set which is disconnected by the omission of any two of its points.

THEOREM 4. *In a space satisfying Axioms* $1'$ *and* 4 *of F.A., a set of points* M *satisfying the requirements of Definition* 4 *is the sum of two simple continuous arcs that have only their endpoints in common.*

Proof. Suppose the set of points M satisfies the requirements of Definition 4. Let A and B denote two distinct points of M. By hypothesis the set $M - (A + B)$ is the sum of two mutually separated point-sets M_1 and M_2. I will show that $M_1 + A + B$ and $M_2 + A + B$ are simple continuous arcs from A to B. That $M_1 + A + B$ is closed is evident.

It is also connected. For suppose it is not. Then it is the sum of two closed, mutually exclusive point-sets N and K. If N should contain both A and B then M would be the sum of the two separated sets $N + M_2$ and K which is contrary to hypothesis. Similarly K cannot contain both A and B. It follows that one of the sets N and K contains A and the other contains B. Suppose K contains A. The set $M_2 + A + B$ must be connected. For if

* J. R. Kline, *The converse of the theorem concerning the division of a plane by an open curve*, these Transactions, vol. 18 (1917), pp. 177–184.

it were the sum of two mutually separated sets $\overline{N}$ and $\overline{K}$ where $\overline{K}$ contains both A and B then M would be the sum of the two mutually separated sets $\overline{K} + M_1$ and $\overline{N}$, while if it were the sum of two mutually separated sets $\overline{N}$ and $\overline{K}$ where $\overline{N}$ contains B and $\overline{K}$ contains A then M would be the sum of the two mutually separated sets $K + \overline{K}$ and $N + \overline{N}$. The set K is connected. For otherwise it would be the sum of two mutually separated sets K_1 and K_2 (where K_1 contains A) and M would be the sum of two separated sets $M_2 + K_1 + N$ and K_2. Likewise N is connected. Hence $N + M_2 + A$ is connected. Either K or N contains more than one point.

Case I. Suppose that K contains more than one point but N contains only the point B and that the set $M_2 + A$ is not connected. Then $M_2 + A$ is the sum of two mutually separated point-sets L_1 and L_2 where L_2 contains A, B is a limit point both of L_1 and of L_2, the set $M - B$ is the sum of the two mutually separated sets L_1 and $L_2 + K$, and the set $M - (A + B)$ is the sum of the two separated sets $L_2 - A$ and $K - A + L_1$. That this leads to a contradiction may be proved by an argument entirely analogous to that employed in the sub-case of Case II below in which N contains more than one point.*

Case II. Suppose that K contains more than one point and that either $M_2 + A$ is connected or N contains more than one point. In the first case let Y_0 denote the point B. Otherwise let Y_0 denote some definite point of N other than B. In either case, if X is any point of K other than A, $M - X - Y_0$ is the sum of two mutually separated sets M_{XY_0} and $\overline{M}_{XY_0}$ where M_{XY_0} contains $M_2 + A$. The set $\overline{M}_{XY_0} + X + Y_0$ is the sum of two sets K_X and N_{Y_0} where K_X is a subset of K and N_{Y_0} is a subset of N. For every point X of K distinct from A, the set K_X is connected. For otherwise it would be the sum of two separated sets K_{X1} and K_{X2} where K_{X1} contains X and in this case the set M would be the sum of two mutually separated sets, K_{X2} and $N + M_{XY_0} + K_{X1}$, which is contrary to hypothesis. The set $M_{XY_0} + X + Y_0 + N_{Y_0}$ is connected. For if it were the sum of two mutually separated sets L and T where L contains X then M would be the sum of two mutually separated sets, $L + K_X$ and T, which is contrary to hypothesis. Suppose that X_1 and X_2 are two points of K and that X_2 is in K_{X_1}. If X_1 were in K_{X_2} then the connected point-set $M_{X_2Y_0} + X_2 + Y_0 + N_{Y_0}$ would contain one point X_2 of K_{X_1} but would not contain the point X_1 and therefore would necessarily be a subset of K_{X_1} and therefore of K which is not the case. Hence X_1 is not in K_{X_2}. But K_{X_2} is connected and, by hypothesis, contains a point X_2 in K_{X_1}. Hence K_{X_2} is a subset of K_{X_1}. Suppose now that X_1 and X_2 are two points in K and that K_{X_1} and K_{X_2} have a point in common. Then, unless X_2 is in K_{X_1}, K_{X_2} must contain the whole of K_{X_1}. But if X_2 is in

*In the proof in question below merely replace N by L_1 and M_2 by $L_2 - A$.

K_{X_1} then, as has been shown above, K_{X_1} contains the whole of K_{X_2}. We thus have the result that if K_{X_1} and K_{X_2} have a point in common then one of them contains the other one. The points of K can be arranged in a well-ordered sequence β. Let P_1 be the first point in the sequence β. Let P_1 be the first point of β which lies in K_{P_1}. Let P_3 be the first point of β which is common to K_{P_1} and K_{P_2}. Let P_3 be the first point of β which is common to K_{P_1}, K_{P_2}, and K_{P_3}. This process may be continued. It follows that the sequence β contains a subsequence α such that if T is a subset of the elements of α there is an element of α which follows, in α, all the elements of T if, and only if, there exists a point which belongs to K_P for every point P of T and, if there does exist such a point, then the first element of α that follows all the elements of T is the first point of β which belongs to K_P for every point P of T. For every two distinct points X and Y of the sequence α either K_X contains K_Y or K_Y contains K_X. Moreover the set K is bounded. It follows* that there exists a point P which belongs to every K_X for every point X of α. It is clear that there is only one such point P and that K_P consists of the single point P. For each point X of α the set $M_{XY_0} + X + Y_0$ is connected. But every point of $M - P$ belongs to $M_{XY_0} + X + Y_0 + N_{Y_0}$ for some point X of α. It follows that both $M_{PY_0} + Y_0$ and $M - P$ are connected. If Y is any point of N_{Y_0} then $\overline{M}_{PY} + P + Y = P + N_Y$. By an argument similar to that used to establish the existence of P it can be shown that there exists in N a point Q such that $N_Q = Q$. The set $M - (P + Q)$ is connected. Thus the supposition that $M_1 + A + B$ is not connected has led to a contradiction. Similarly $M_2 + A + B$ is connected.

Suppose now that P_1 is any point of M_1. Let P_2 denote any point of M_2. The set $M - (P_1 + P_2)$ is the sum of two mutually separated sets $\overline{M}_1$ and $\overline{M}_2$. Since $\overline{M}_1 + P_1 + P_2$ and $\overline{M}_2 + P_1 + P_2$ are connected, P_1 is a limit point both of $\overline{M}_1$ and of $\overline{M}_2$. It follows that both of these sets contain points of M_1. Let $\overline{N}_1$ denote the set of points common to $\overline{M}_1$ and $M_1 + A + B$ and let $\overline{N}_2$ denote the set of points common to $\overline{M}_2$ and $M_1 + A + B$. The set $(M_1 + A + B) - P_1$ is the sum of the two mutually separated sets $\overline{N}_1$ and $\overline{N}_2$.

It has now been shown that the point-set $M_1 + A + B$ satisfies all the requirements of a simple continuous arc from A to B as given in Definition 2. Similarly $M_2 + A + B$ is a simple continuous arc from A to B. Clearly these two arcs have only A and B in common. The truth of Theorem 4 is therefore established.

* See my paper, *On the most general class L of Fréchet in which the Heine-Borel-Lebesgue theorem holds true*, loc. cit.

3

DEFINITION. If the point-set M is a proper subset of the point-set N, the *boundary* of M with respect to N is the set of all points $[X]$ such that X is either a point or a limit point of M and also either a point or a limit point of $N - M$.

THEOREM 5. *If the continuous* point-set M contains no continuous set of condensation† then every two points of M are the extremities of a simple continuous arc that lies wholly in M.*

Indication of proof. By an argument largely similar to (but not entirely identical with) one used in my paper, *Concerning continuous sets that have no continuous sets of condensation*,‡ it may be proved that M is "connected in kleinem." By an argument similar to that used in my paper, *A theorem concerning continuous curves*,§ it may be proved that every two points of M are the extremities of a simple continuous arc lying wholly in M.

THEOREM 6. *In euclidean space of two dimensions if no continuous subset of the continuous point-set M has more than two boundary points with respect to M then M is a simple continuous arc, a simple closed curve, a simple open curve, or a ray of a simple open curve.*

Proof. The set M contains no continuous set of condensation. Hence by Theorem 5 every two points of M are the extremities of at least one simple continuous arc that lies wholly in M. Let G denote the family of all arcs which lie in M. If AB is an arc of G no point of AB distinct from A and from B is a limit point of $M - AB$. For if there should exist such a point P there would exist on AB two points $\overline{A}$ and $\overline{B}$ in the order $A\overline{A}P\overline{B}B$ on AB and the interval $\overline{A}\overline{B}$ would be a continuous subset of M that has three distinct boundary points with respect to M, these boundary points being the points $\overline{A}$, P, and $\overline{B}$.

If g_1 and g_2 are two arcs of G with a common point then the point-set $g_1 + g_2$ is either a simple continuous arc or a simple closed curve. For otherwise one of the arcs g_1 and g_2 (call it g) would contain a point P which is not an endpoint of g but which is a limit point of $M - g$.

Let A and B denote two definite distinct points of M and let AB denote a definite arc of the set G having A and B as endpoints. Let K_C $(C = A, B)$ denote the set of all points X of M such that C and X are the extremities of a simple continuous arc which is a subset of M but which contains no point except C in common with AB. An arc from C to X satisfying these conditions

* A set of points is said to be *continuous* if it is closed and connected and contains more than one point.

† The continuous set of points N is said to be a *continuous set of condensation* of the set M if N is a proper subset of M and every point of N is a limit point of $N - M$.

‡ Bulletin of the American Mathematical Society, vol. 25 (1919), pp. 174–176.

§ Ibid., vol. 23 (1917), pp. 233–236.

will be called the arc CX. If neither K_A nor K_B exists then M is the arc AB. Suppose that K_A exists. If X_1 and X_2 are points of K_A, X_1 is said to precede X_2 if AX_1 is a subset of AX_2. It is clear that of any two distinct points of K_A one of them precedes the other one. The set $K_A + A + B$ is closed. For suppose there exist two points O and $\overline{O}$, distinct from A, which are limit points of K_A but which do not belong to K_A. It can easily be proved with the help of Theorem 5 of F.A. that there exist two sets of points $A_1, A_2, A_3, \cdots$ and $\overline{A}_1, \overline{A}_2, \overline{A}_3, \cdots$ belonging to K_A such that, for every n, A_n precedes $\overline{A}_n$ and $\overline{A}_n$ precedes A_{n+1} and such that O is the sequential limit point of $A_1, A_2, A_3, \cdots$ while $\overline{O}$ is the sequential limit point of $\overline{A}_1, \overline{A}_2, \overline{A}_3, \cdots$. It follows, by an argument similar to that used in establishing the existence of the set t in the proof of Theorem 1, that there exists a continuous set of points $\overline{K}$ containing O such that if Y is a point of $\overline{K}$ then every region about Y contains points of infinitely many of the arcs $A_1A_2, A_2A_3^*, \cdots$. For no value of n does the arc A_nA_{n+1} contain a limit point of the point-set composed of the arcs $A_{n+2}A_{n+3}, A_{n+3}A_{n+4}, \cdots$. Hence $\overline{K}$ contains no point of any of the arcs $A_1A_2, A_2A_3, \cdots$. It follows that $\overline{K}$ is a continuous set of condensation of M. But it was shown above that M contains no continuous set of condensation. Thus the supposition that there exist two limit points of $K_A + A$ that do not belong to $K_A + A$ has led to a contradiction. It follows that either $K_A + A + B$ is closed or there exists one and only one point P, distinct from B, which does not belong to K_A but is a limit point of K_A. In the latter case P is a sequential limit point of a sequence of points $P_1, P_2, P_3, \cdots$ belonging to K_A such that, for every n, P_n precedes P_{n+1} and the point-set $P + A_1P + P_1P_2 + P_2P_3\dagger + \cdots$ is a simple continuous arc from A to P. It follows that P either coincides with B or belongs to K_A. But this is contrary to supposition. Hence the set $K_A + A + B$ is closed.

There are several cases to be considered.

Case I. Suppose that $K_A + A$ is unbounded and that K_B does not exist. The set $K_A + A$ contains a countably infinite set of distinct points $A_1, A_2, A_3, \cdots$ such that $A_1 + A_2 + A_3 + \cdots$ has no limit point. Let $\overline{A}_2$ be the first point of this sequence that follows A_1, $\overline{A}_3$ the first one that follows $\overline{A}_2$, etc. There results an infinite sequence $\overline{A}_1, \overline{A}_2, \overline{A}_3, \cdots$ of points belonging to the set $A_1, A_2, A_3, \cdots$ such that, for every n, $\overline{A}_n$ precedes $\overline{A}_{n+1}$. For each n there exists a region R_{n+1} containing $\overline{A}_{n+1}$ but containing no point of M that does not belong to the interval $\overline{A}_n\overline{A}_{n+2}$ of M. It can be shown that there exists a countable set of distinct points $\overline{B}_1, \overline{B}_2, \overline{B}_3, \cdots$ (where $\overline{B}_1 = B$), with no limit point, and a set of arcs $\overline{B}_1\overline{B}_2, \overline{B}_2\overline{B}_3, \overline{B}_3\overline{B}_4, \cdots$ such that

* Here, for every n, A_nA_{n+1} denotes the interval of the arc AA_n whose endpoints are A_n and A_{n+1}.

† Here, for every n, P_nP_{n+1} denotes the interval of AP_{n+1} whose endpoints are P_n and P_{n+1}.

(1) for every n, $\overline{B}_{n+1}$ is in R_{n+1}, (2) the set $M + \overline{B}_1\overline{B}_2 + \overline{B}_2\overline{B}_3 + \cdots$ is an open continuous curve. The set M is a ray of this curve.

Case II. Suppose that K_A is unbounded and that K_B exists but is bounded. Then K_A and K_B can have no point in common and the set K_B cannot contain more than one last point. If P is a point of K_B which is not a last point then $K_B + B - P$ is not connected. Thus the set $K_B + B$ is a simple continuous arc (see Definition 2). That $(K_B + B) + K_A$ is a ray follows by an argument similar to that employed in Case I.

Case III. Suppose K_A and K_B are both unbounded. In this case the closed sets K_A and K_B have no point in common and the set $K_B + \overline{AB} + K_A$ is evidently an open curve.

Case IV. Suppose K_A is bounded and that K_B either is bounded or does not exist and that K_A and K_B have no point in common. In this case M is clearly a simple continuous arc.

Case V. Suppose that K_A is bounded and that K_B and K_A have at least one point in common. It is clear that in this case $K_A + B$ is a simple continuous arc from A to B and that M is a simple closed curve.

If the term simple continuous curve is applied only to point-sets which are either arcs, closed curves, open curves, or rays then it is clear in view of the above results that these point-sets may be defined as follows.

DEFINITION 5. A *simple continuous curve** is a continuous point-set M no continuous subset of which has more than two boundary points with respect to M.

DEFINITION 6. A *simple closed curve* is a continuous point-set M every proper continuous subset of which has just two boundary points with respect to M.

DEFINITION 7. A *simple continuous open curve* is a continuous point-set M every bounded continuous subset of which has just two boundary points with respect to M.

DEFINITION 8. A *simple continuous arc* is a continuous curve M containing a point A such that every continuous proper subset of M that contains A has just one boundary point with respect to M.

DEFINITION 9. A *ray* is a continuous curve M containing a point A such that every bounded continuous subset of M that contains A has just one boundary point with respect to M.

* The term continuous curve is ordinarily not applied to sets that are not bounded. Thus, according to the ordinary terminology, a straight line is not a continuous curve. I suggest that the term continuous curve be applied to every closed point-set which is connected "in kleinem" (in the sense of H. Hahn) whether it be bounded or unbounded. If this terminology is adopted what is now ordinarily called a continuous curve may be characterized as a *bounded continuous curve*.

UNIVERSITY OF TEXAS

ON THE ITERATION OF RATIONAL FUNCTIONS*

BY

J. F. RITT

I. Introduction

Notes on the iteration of rational functions have been published in the Comptes Rendus during the last few years by Julia, Fatou, Lattés, and myself.† Many of the results announced were obtained independently by more than one writer, the priority in such cases seeming to belong to Julia. I wish to present here that small part of my work which has not been covered by the other writers, postponing until after the appearance of their memoirs the publication of such of my other results as may still be of interest.‡

The pth iterate of the analytic function $f(x)$ will be denoted by $f_p(x)$. If $f(c) = c$, the point c will be called a *fixed point* of $f(x)$. Let the Taylor development of $f(x)$ at the fixed point c be

$$f(x) = c + a_1(x-c) + a_2(x-c)^2 + \cdots + a_n(x-c)^n + \cdots.$$

The point c will be called a *point of attraction* or *point of repulsion* according as $|a_1| < 1$ or $|a_1| > 1$.§ I have not had occasion thus far to study the case where $|a_1| = 1$. If $\alpha_1 = f_p(\alpha_2)$, where p is some positive integer, α_1 is called a *consequent* of α_2, and α_2 an *antecedent* of α_1. If $f_p(\alpha_1) = f_p(\alpha_2)$ for some positive integer p, I shall say that α_1 and α_2 are *associates* of each other. If

$$F(x) = \phi^{-1} f \phi(x),$$

$f(x)$ will be said to be transformed into $F(x)$ by means of $\phi(x)$.

I have made use, in my work, of a theorem which was first found by L.

* Presented to the Society, December 27, 1917, and February 23, 1918.

† Julia, December 31, 1917, January 14, 1918, January 28, 1918; Fatou, May 14, 1917, December 17, 1917, February 4, 1918; Lattés, January 7, 1918, January 28, 1918, March 25, 1918; Ritt, March 4, 1918.

‡ While this paper was in the hands of the editors of the Transactions, papers appeared by Julia, Journal de Mathématiques, 1918 (published one year late), and by Fatou, Bulletin de la Société Mathématique de France, 1919. Fatou's paper contains some of the details of § III of the present paper, but as far as I can see, these results were first stated in my note in the Comptes Rendus.

§ These terms were introduced by me in my note in the Comptes Rendus, and have since been used by all other writers on this subject.

Boettcher,* who published it in Russian. To make the present paper easier to read, I have devoted § II to a very brief proof of Boettcher's theorem.

In § III I discuss the set of associates of a point.

In § IV I determine the conditions under which two polynomials can have an iterate in common.

II. Boettcher's theorem

Let $f(x)$ have, for $x = 0$, the Taylor development

$$f(x) = a_m x^m + a_{m+1} x^{m+1} + \cdots + a_{m+i} x^{m+i} + \cdots,$$

where $m > 1$. Within a sufficiently small circle with center at the origin, $f_p(x)$ approaches zero uniformly as p increases indefinitely. We consider a circle of this type, in which $f(x)$ has no zero other than $x = 0$. Within this circle, every iterate $f_p(x)$, $(p = 1, 2, \cdots, n, \cdots)$, exists, and vanishes only for $x = 0$, at which point it has a zero of order m^p. The m^pth roots of $f(x)$ will be uniform within this circle, with simple zeros at $x = 0$. The first coefficient in the development of $f_p(x)$ is $a_m^{(m^p-1)/(m-1)}$. We can thus select a sequence of functions

$$(1) \qquad [f(x)]^{1/m}, \quad [f_2(x)]^{1/m^2}, \quad \cdots, \quad [f_p(x)]^{1/m^p}, \quad \cdots,$$

such that the coefficients of x in the terms of the sequence approach any chosen $(m-1)$th root of a_m. We may now state

Boettcher's Theorem. *In the circle described above, the functions in the sequence* (1) *converge uniformly to an analytic function* $\phi(x)$, *which satisfies the functional equation*

$$(2) \qquad \phi f(x) = [\phi(x)]^m.$$

We suppose first that $a_m = 1$; the general case will be shown later to depend on this special one. We may assume then that the coefficients of x in the terms of (1) are all unity.

The functions of the sequence (1) have a common upper bound for their moduli, and hence also a common majorant, in the circle under consideration. We may thus write, since as p increases, $f_p(x)$ approaches zero uniformly with respect to x,

$$f_{p+q}(x) = f_q[f_p(x)] = [f_p(x)]^{m^q}[1 + \epsilon_{p,q}(x)],$$

where $\epsilon_{p,q}(x)$ tends toward zero as p increases, uniformly with respect to x and q. Consequently,

$$(3) \qquad [f_{p+q}(x)]^{1/m^{p+q}} = [f_p(x)]^{1/m^p}[1 + \eta_{p,q}(x)],$$

*Bulletin of the Kasan Mathematical Society, vol. 14 (1905), p. 176.

where $\eta_{p,q}(x)$ approaches zero as x increases, uniformly with respect to x and q. The terms of the sequence (1) are, after a certain p, less than unity in modulus, since they are roots of functions which approach zero. It follows from (3) that the difference between the $(p+q)$th and the pth terms of the sequence (1) approaches zero as p increases, uniformly with respect to x and q. This proves that the functions of the sequence (1) converge uniformly to an analytic function $\phi(x)$. Furthermore, since

$$\{f_p[f(x)]\}^{1/m^p} = \{[f_{p+1}(x)]^{1/m^{p+1}}\}^m,$$

we find, on allowing p to approach infinity,

$$\phi f(x) = [\phi(x)]^m. \tag{2}$$

As to the case where a_m is not unity, let

$$F(x) = a_m^{1/(m-1)} f(x/a_m^{1/(m-1)}).$$

The coefficient of x^m in $F(x)$ is unity. Let

$$\Phi(x) = \lim_{p\to\infty} [F_p(x)]^{1/m^p}.$$

Since

$$[f_p(x)]^{1/m^p} = a_m^{-1/m^p(m-1)} [F_p(a_m^{1/(m-1)}x)]^{1/m^p},$$

it is clear that the functions of the sequence (1) approach as a limit

$$\phi(x) = \Phi(a_m^{1/(m-1)}x),$$

and that $\phi(x)$ satisfies equation (2). This concludes the proof of Boettcher's theorem.

If $\phi(x)$ satisfies the functional equation (2), the function $\epsilon[\phi(x)]^i$, where i is any integer, positive or negative, and ϵ any $(m-1)$th root of unity, also satisfies the equation. Furthermore, this is the most general solution of equation (2) which is meromorphic in the neighborhood of the origin, for $f(x)$ being given, and the degree of the first term in the development of $\phi(x)$ having been assigned, equation (2) determines the coefficient of the first term of $\phi(x)$ as one of the $(m-1)$th roots of a_m, and determines the other coefficients uniquely.

If $f(x)$ is an analytic function with a pole of order m at infinity, we can transform it into a function with a zero of order m at the origin by means of the function $1/x$. Simultaneously, the functions of the sequence (1), calculated for $f(x)$, will be transformed into the corresponding functions calculated for $1/f(1/x)$. Hence the functions of (1) approach a function $\phi(x)$ with a simple pole at infinity, which satisfies equation (2). This fact will be used in § IV.

III. The associated points

Let $f(x)$ be a rational function of degree greater than unity.

If $f(x)$ is a polynomial, the point infinity has no associates other than itself. If $f(x)$ is a positive or negative power of x, the points zero and infinity each have no associates different from themselves. If $f(x)$ can be transformed into a polynomial, or into a power of x, by means of a linear function, integral or fractional, there are one or two points, respectively, which have no associates different from themselves.

We shall show that except in the cases described above, every point has an infinite number of associates.

Suppose that a point c has only a finite number of associates. Let $f_\nu(c)$ be one of the consequents which c and all of its associates have in common. Then $f_\nu(c)$ can have no associates other than itself, for the νth antecedents of such associates would be associates of c, and would be distinct from the νth antecedents of $f_\nu(c)$.

Consider first the case where $f_\nu(c)$ is not a fixed point of $f(x)$. The points $f_\nu(c)$ and $f_{\nu+1}(c)$ have no associates different from themselves. Take a linear function $\phi(x)$, such that $\phi^{-1}f_\nu(c) = 0$, and $\phi^{-1}f_{\nu+1}(c) = \infty$. Then, relative to the rational function $\phi^{-1}f\phi(x)$, the points zero and infinity, and their consequents, have no associates different from themselves. We must consequently have

$$\phi^{-1}f\phi(x) = \frac{a + bx^n}{x^n},$$

where $n > 1$. We have $\phi^{-1}f\phi(\infty) = b$. If b were not zero, it would have associates different from itself, which is impossible. Consequently $f(x)$ can be transformed into the form a/x^n, and hence also into $1/x^n$.*

Consider now the case where $f_\nu(c)$ is a fixed point of $f(x)$. Then $f_\nu(c)$ can have no antecedent different from itself, for such antecedents would also be its associates. Let $f(x)$ be transformed with a linear function $\phi(x)$ such that $\phi^{-1}f_\nu(c) = \infty$. Then, relative to $\phi^{-1}f\phi(x)$, the point infinity has no antecedent different from itself. The function $\phi^{-1}f\phi(x)$ is consequently a rational integral function.

If there is a point d, distinct from $f_\nu(c)$, which has only a finite number of associates relative to $f(x)$, a common consequent $f_\mu(d)$ of d and its associates must be a fixed point of $f(x)$, for it is easily seen that a function of degree higher than unity cannot be transformed both into a rational integral function and into a negative power of x. If $\phi^{-1}f_\mu(d) = 0$, and $\phi^{-1}f_\nu(c) = \infty$, the function $\phi^{-1}f\phi(x)$ will be of the form cx^n. Hence $f(x)$ can be transformed into the form x^n.

* The function a/x^n is transformed into $1/x^n$ by means of $\phi(x) = a^{1/n+1}/x$.

The question as to the number of associates of a point being thus settled, we consider a point c, which has an infinite number of associates relative to the rational function $f(x)$. The following facts will be proved:

I. *The set of limit points of the associates of c contains a perfect set.*

II. *The associates of c cannot accumulate about a point of attraction.*

III. *The consequents of a limit point of the associates of c cannot approach a point of attraction at which the derivative of $f(x)$ is not zero, even should the consequents of c approach the point of attraction.*

IV. *If the consequents of c tend toward a point of attraction at which the derivative of $f(x)$ is not zero, the associates of c form an isolated set.*

V. *If c lies within a sufficiently small neighborhood of a point of attraction at which the derivative of $f(x)$ is zero, there exists a simple closed analytic curve, passing through c, and around the point of attraction, on which the associates of c are everywhere dense. If $f(x)$ has a pole of order greater than one at infinity, the associates of a point in the neighborhood of infinity lie everywhere dense on a simple closed analytic curve which passes through the point.*

We shall prove the result V first. Let the point of attraction be taken as the point $x = 0$. This can always be brought about by transforming $f(x)$ with a function of the form $x + h$. Let the first term in the development of $f(x)$ about $x = 0$ be of degree m, where $m > 1$. By Boettcher's Theorem there exists an analytic $\phi(x)$, with a simple zero at $x = 0$, such that

$$f(x) = \phi^{-1}\{[\phi(x)]^m\}.$$

The associates of any point, relative to the function x^m, lie everywhere dense on a circle passing through the point, with center at the origin. If d is an associate of c, relative to $f(x)$, $\phi(d)$ will be an associate of $\phi(c)$, relative to the function x^m. For sufficiently small values of c, the function $\phi^{-1}(x)$ transforms the circle passing through $\phi(c)$ and with center at the origin, into a simple closed analytic curve, passing through c and around the origin. The associates of c, relative to $f(x)$, are everywhere dense on this curve. This proves the first statement in V. The second statement is proved by transforming $f(x)$ with the function $1/x$.

In the first case of V, if the degree of $f(x)$ is greater than m, the order of the zero of $f(x)$ at $x = 0$, c will have associates which are not included among the associates which were shown above to lie on a curve about the origin. If c is sufficiently small, none of these additional associates can lie within the closed curve, else the point $\phi(c)$ would have an associate relative to the function x^m lying within the circle described above. It follows that the associates of a point cannot accumulate about a point of attraction at which the derivative of $f(x)$ is zero. We shall show that the same holds if the derivative of $f(x)$ is not zero, thus completing the proof of II. Let c and d be two points very

close to the point of attraction which have a common consequent. The consequents of c and d must approach the point of attraction. Let ν be the smallest integer such that $f_\nu(c) = f_\nu(d)$. Then the points $f_{\nu-1}(c)$ and $f_{\nu-1}(d)$, which are very close to the origin, are distinct, and the function $f(x)$ has the same value at both of these points. This is impossible, since the derivative of $f(x)$ does not vanish at the point of attraction. The proof of II is completed.*

Result III follows immediately from the argument given above, for if the consequents of a limit point of the associates of c should tend toward a point of attraction at which the derivative of $f(x)$ is not zero, there would be distinct points in every neighborhood of this point of attraction which have a common consequent. Result IV is a corollary of III.

We shall now prove I. If d is a limit point of the associates of c, all of the associates of d are likewise limit points of the associates of c. Such a point d cannot be an exceptional point which has only a finite number of associates, since those special points are points of attraction. Thus D_1, the first derived assemblage of the limit points of the associates of c contains an infinite number of points. Hence the first derived assemblage D_2, of D_1, contains at least one point. It is seen readily that if a point belongs to D_2, all of its associates also belong to D_2. Hence D_2 contains an infinite number of points. We see that the derived assemblages

$$D_1, D_2, \cdots, D_n, \cdots,$$

all contain an infinite number of points. There exists a closed set of points D_ω common to the sets of the above sequence. The associates of every point of D_ω belong to D_ω, so that D_ω contains an infinite number of points. Continuing thus, we can demonstrate the existence of derived assemblages of every transfinite order. It follows from a well-known theorem that the set of limit points of c contains a perfect set.†

IV. Polynomials with a common iterate

We shall discuss in this section the circumstances under which two rational integral functions can have an iterate in common. The discussion given does not apply to the case where the functions are linear, which trivial case can be handled by a direct comparison of the iterates. Also, to avoid the introduction of extra details, we shall bar out the case where the functions can be transformed into multiples of powers of x, in which case the iterates can also be compared directly.

* If, on transforming $f(x)$ so that ∞ is carried into a finite point, this finite point is a point of attraction of the transformed function, we may say that ∞ is a point of attraction of $f(x)$. In that case the associates of a point cannot accumulate about ∞.

† This argument holds even if the set of associated points is not bounded.

Let us suppose then that the two non-linear polynomials, $f(x)$ and $F(x)$, of degrees n and N respectively, are such that

$$f_p(x) = F_P(x).^*$$

We must have

$$n^p = N^P.$$

There exist therefore three integers, r, s, and S, such that

$$n = r^s, \qquad N = r^S.$$

Consider the two Boettcher functions,

$$\phi(x) = \lim_{i\to\infty} [f_i(x)]^{1/n^i}, \qquad \Phi(x) = \lim_{i\to\infty} [F_i(x)]^{1/N^i},$$

which exist in the neighborhood of ∞, at which point they have simple poles. Since

$$f_{ip}(x) = F_{iP}(x)$$

for every integer i, we must have

$$\Phi(x) = c\phi(x),$$

where c is a constant of modulus unity. Then

$$f(x) = \phi^{-1}\{[\phi(x)]^r\}, \qquad F(x) = \phi^{-1}\{l[\phi(x)]^{r^S}\},$$

where

$$l = c^{r^S-1}.$$

The lemma which follows will be important for our purposes.

LEMMA. *If the functions*

$$U(x) = \phi^{-1}\{h[\phi(x)]^u\}, \qquad V(x) = \phi^{-1}\{k[\phi(x)]^v\},$$

where $u \geqq v$, *are rational and integral, the function*

$$UV^{-1}(x) = \phi^{-1}\{hk^{1/r^s}[\phi(x)]^{u-v}\},$$

where any of the determinations of $V^{-1}(x)$ *is taken, is also rational and integral.*

From the expression of $UV^{-1}(x)$ in terms of $\phi(x)$, we see that $UV^{-1}(x)$ is uniform in the neighborhood of infinity. Let its Laurent development be

$$G(x) + P(1/x),$$

where $G(x)$ is a rational integral function, and $P(1/x)$ a series of negative powers of x. We have

$$U(x) = G[V(x)] + P[1/V(x)].$$

The Laurent development of $P[1/V(x)]$ for the point infinity can contain only negative powers of x, and since the function $U(x)$ has no negative powers

* I solved this problem originally for the case of $p = P$. Dr. Gronwall suggested to me the possibility of generalization.

in its expansion about infinity, $P[1/V(x)]$, and hence also $P[1/x]$, must be identically zero. Therefore

$$UV^{-1}(x) = G(x),$$

which proves the lemma.

Let t be the smallest integer greater than zero such that there exists a rational integral function of the form

$$g(x) = \phi^{-1}\{e[\phi(x)]^{r^t}\}.$$

Then t must be a divisor of s. For otherwise we would have

$$s = it + j,$$

where i and j are integers and $0 < j < t$, and the function $f[g_{-i}(x)]$, $g_{-i}(x)$ being the inverse of the ith iterate of $g(x)$, would be a rational integral function of the form

$$\phi^{-1}\{b[\phi(x)]^{r^j}\}.$$

Similarly t is a divisor of S.

If a polynomial whose Boettcher function is $\phi(x)$ is transformed with a linear integral function $\omega(x)$, the Boettcher function of the transformed polynomial will be $\phi[\omega(x)]$. We may therefore suppose that the functions $f(x)$ and $F(x)$ have been transformed with a function of the form $x + h$ so that the development about infinity of the Boettcher function $\phi(x)$ contains no constant term.

Putting $q = s/t$, we see from the lemma above that $f(x)$ is a linear function of $g_q(x)$. Since this linear function is of the form

$$\phi^{-1}[h\phi(x)],$$

it cannot contain a constant term. We have thus

$$f(x) = \epsilon_1 g_q(x),$$

where ϵ_1 is a constant. Similarly

$$F(x) = \epsilon_2 g_Q(x),$$

where ϵ_2 is a constant and where $Q = S/t$.

Expressing $f(x)$ and $g_q(x)$ in terms of $\phi(x)$, we have

$$\phi^{-1}\{[\phi(x)]^{r^s}\} = \epsilon_1 \phi^{-1}\{\gamma[\phi(x)]^{r^t}\},$$

where

$$\gamma = e^{(r^{qt}-1)/(r^t-1)}.$$

Hence

$$\phi^{-1}(x) = \epsilon_1 \phi^{-1}(\gamma x).$$

Let

$$\phi^{-1}(x) = \alpha x + \frac{\alpha_1}{x} + \frac{\alpha_2}{x^2} + \cdots + \frac{\alpha_n}{x^n} + \cdots.$$

Since $f(x)$ is not a multiple of a power of x, $\phi^{-1}(x)$ must actually contain

negative powers. One sees easily from this expression for $\phi^{-1}(x)$ that $\epsilon_1 = \gamma$ and that ϵ_1 is a root of unity. We must have

$$\phi^{-1}(x) = \alpha x P(1/x^m),$$

where $\epsilon_1^m = 1$, and where $P(1/x^m)$ is a series of powers of $1/x^m$. Hence, also,

$$\phi(x) = \frac{x}{\alpha} Q(1/x^m),$$

where $Q(1/x^m)$ is a series of negative powers of x^m.

We see now from the expression for $g(x)$ in terms of $\phi(x)$ that

$$g(x) = x^{r'} R(1/x^m),$$

where $R(1/x^m)$ is a polynomial in $1/x^m$.

If we express $F(x)$ and $g_Q(x)$ in terms of $\phi(x)$, we have

$$\phi^{-1}\{l[\phi(x)]^{r^S} = \epsilon_2 \phi^{-1}\{\delta[\phi(x)]^{r^S}\},$$

where δ is easily expressed as a power of l. We see as above that

$$\epsilon_2 = \frac{\delta}{l}$$

and that ϵ_2 is a root of unity. Also, if the integer m introduced above is taken as large as possible, so that there are actually two terms in $P(1/x^m)$ whose exponents differ by m, we have $\epsilon_2^m = 1$.

We have

$$g_q(\epsilon_1 x) = \epsilon_1^{r^q} g_q(x), \qquad g_Q(\epsilon_2 x) = \epsilon_2^{r^S} g_Q(x).$$

Consequently,

$$f_p(x) = \epsilon_1^{(r^{pq}-1)/(r^q-1)} g_{pq}(x) = \epsilon_1^{(n^p-1)/(n-1)} g_{pq}(x),$$

$$F_P(x) = \epsilon_2^{(r^{PS}-1)/(r^S-1)} g_{PQ}(x) = \epsilon_2^{(N^P-1)/(N-1)} g_{PQ}(x).$$

Therefore

$$\epsilon_1^{(n^p-1)/(n-1)} = \epsilon_2^{(N^P-1)/(N-1)}.$$

The conditions on $f(x)$ and $F(x)$ found thus far are easily seen to be sufficient. Our result may be stated as follows: *There exists a polynomial $g(x)$, in which the exponents form an arithmetic progression with difference m, such that $f(x)$ and $F(x)$ can be derived by transforming*

$$\epsilon_1 g_q(x), \qquad \epsilon_2 g_Q(x)$$

respectively, with a function of the form $x + h$. The constants ϵ_1 and ϵ_2 satisfy the relations

$$\epsilon_1^m = 1, \qquad \epsilon_2^m = 1,$$

$$\epsilon_1^{(n^p-1)/(n-1)} = \epsilon_2^{(N^P-1)/(N-1)}.$$ *

* It should be remembered that $n^p = N^P$.

COLUMBIA UNIVERSITY